

Irina Pinter

Stochastische Modellierung von Preisprozessen mit
Hilfe von Mischungsverteilungshypothese

Masterarbeit

zur Erlangung des akademischen Grades
einer Magistra der Sozial- und Wirtschaftswissenschaften
der Studienrichtung Betriebswirtschaft
an der Karl-Franzens-Universität Graz

Ao. Univ.-Prof. Dr. Roland Mestel

Institut für Banken und Finanzierung

Graz, August 2009

Ehrenwörtliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen inländischen oder ausländischen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Die vorliegende Fassung entspricht der eingereichten elektronischen Version.

Datum:

Unterschrift:

Inhaltsverzeichnis

Abbildungsverzeichnis	ii
Tabellenverzeichnis	ii
Abkürzungs- und Symbolverzeichnis	iii
1. Einführung	1
2. Theoretische Grundlagen	4
2.1 Einführung	4
2.2 Statistische Eigenschaften der Renditezeitreihen.....	4
2.3 Statistische Eigenschaften der Handelsvolumenzeitreihen.....	5
2.4 <i>Short memory</i> und <i>long memory</i> Prozesse	6
2.5 Mischungsverteilungen	9
2.6 Zusammenfassung.....	11
3. Entwicklung der Mischungsverteilungshypothese	12
3.1 Einführung	12
3.2 Das Standardmodell von Clark	13
3.3 Dynamische Erweiterung des univariaten Modells.....	20
3.3.1 Das ARV-Modell	21
3.3.2 ARCH-Modelle.....	26
3.3.3 Schätzung der Modelle und Vergleich der empirischen Ergebnisse.....	34
3.4 Bivariate Modelle.....	38
3.4.1 Das Ausgangsmodell.....	39
3.4.2 Das modifizierte Modell	44
3.4.3 Schätzung der Modelle und Vergleich der empirischen Ergebnisse.....	48
3.5 Zusammenfassung.....	54
4. Ansätze zur Erklärung des Verhaltens des Persistenzparameters	56
4.1 Einführung	56
4.2 Heterogene Komponenten in der Informationsrate.....	57
4.3 Zeitabhängige Informationssensibilität.....	68
4.4 Unterschiedliche statische und dynamische Effekte	76
4.4.1 Kurzfristiger Zusammenhang	77
4.4.2 Langfristiger Zusammenhang	81
4.5 Zusammenfassung.....	83
5. Zusammenfassung	85
Literaturverzeichnis	86

Abbildungsverzeichnis

Abbildung 1: Logarithmierte Renditen von Andritz AG (2001-2009) im Vergleich zur Normalverteilung.	1
Abbildung 3: Empirische Autokorrelationsfunktion des logarithmierten Handelsvolumens von Siemens, 1994 - 2005.	8
Abbildung 4: Diskretes Beispiel einer Mischungsverteilung.	10
Abbildung 5: Schätzer für den fraktionalen Differenzparameter der quadrierten Renditen und Handelsvolumina ausgewählter Unternehmungen.	83

Tabellenverzeichnis

Tabelle 1: Eigenschaften der Autokorrelationsfunktionen und der spektralen Dichten.	9
Tabelle 2: Das Standardmodell und seine Erweiterungen.	12
Tabelle 3: Schätzungen des Persistenzparameters.	52

Abkürzungs- und Symbolverzeichnis

Lateinische Buchstaben:

B	Backshift-Operator
C	positive Konstante
D	eine nicht weiter spezifizierte Verteilung
D	positive Konstante
E	Erwartungswert
F	Anzahl der Insider
H_0	Nullhypothese
I_{t-1}	Menge der verzögerten Preisänderungen
J	Anzahl der aktiven Händler
K_t	Korrektur aus dem Kalman-Filter
L	Likelihoodfunktion
N	Normalverteilung
N	Anzahl der Intervalle der Diskretisierung
P_i	der aktuelle Marktpreis
P_{ij}^*	Reservationspreis des j -ten Händlers
Po	Poissonverteilung
Q_{ij}	Wunschposition des j -ten Händlers
$\{R_t\}$	renditegenerierender Prozess
UV_t	nicht prognostizierbares Handelsvolumen zum Zeitpunkt t
V_t	Handelsvolumen zum Zeitpunkt t
W	Gewichtungsmatrix
X_T	Matrix der latenten Variablen
Y_t	Variable
Y_T	Matrix der beobachtbaren Variablen
Z	$N(0,1)$ -Zufallsvariable
a	Parameter im ARCH-Modell

a_t	Parameter des verallgemeinerten bivariaten Modells
b	im Abschnitt 3.3: Parameter in den ARCH-Modellen im Abschnitt 4.2: Parameter im bivariaten SV-Modell
c	Konstante
c_t	<i>white noise</i> Innovationen
d	fraktionaler Differenzparameter
e_t	im Abschnitt 3.3: Residuen der Preisänderungen im Abschnitt 4.2: Fehler der Prognose von r_t^2
$f(\omega)$	spektrale Dichte
h_t	bedingte Varianz zum Zeitpunkt t
i	Nummer des Walrasianischen Gleichgewichts
l	Transaktionszeit
$\{l_t\}$	steuernder Prozess
m_0	Rauschkomponente
m_1	Faktor der Sensibilität
m_t	Prognosefehler
\mathbf{m}	Stichprobenmittel
n_t	Informationsrate zum Zeitpunkt t
\bar{n}	positive Konstante
p	Parameter im ARMA-Modell
p_t	logarithmierter Wertpapierkurs zum Zeitpunkt t
q	Parameter im ARMA-Modell
r_t	Preisänderung des Tages t
\mathbf{r}	Vektor der täglichen Preisänderungen
s	im Abschnitt 3.3: stetiger Wertpapierkurs im Abschnitt 4.2: Zustand
t	Kalenderzeit in Tagen
u_t	im Abschnitt 3.2: Innovationen in den Preisänderungen im Abschnitt 3.3: standardnormalverteilte Zufallsvariable im Abschnitt 4.4: Zufallsfehler
v	im Abschnitt 3.2: Mischungsvariable
v_i	Transaktionsvolumen

v_t	Handelsvolumen zum Zeitpunkt t in den Abschnitten 3.3 und 4.4: Innovationen in der bedingten Varianz
v_t^i	Volumenkomponente der Insider
v_t^l	Volumenkomponente der Liquiditätshändler
w	Freiheitsgrade im ARCH- t -Modell
$w(t)$	standardisierte Brownsche Bewegung
w_t	positive Zufallsvariable
$\{y_t\}$	diskreter stochastischer Prozess
z	Innovationen

Griechische und sonstige Buchstaben:

$\Phi(\cdot)$	ARMA-Polynom
$\Gamma(\cdot)$	Gamma-Funktion
Δt	Schrittweite der Diskretisierung
Π	Produktzeichen
$\Theta(\cdot)$	ARMA-Polynom
Σ	Summenzeichen
α	im Abschnitt 2.5: diskrete Mischungsvariable im Abschnitt 3.3: erwartete Momentanrendite des Wertpapiers und ein Parameter in den ARCH-Modellen im Abschnitt 3.4: Konstante
α_0	Parameter in den ARCH-Modellen
α_t	Parameter der Prognose
$\bar{\alpha}$	langfristiger Mittelwert des Prozesses $\ln \sigma^2(t)$
$\tilde{\alpha}_\psi$	Parameter der relativen Wichtigkeit des zeitabhängigen Faktors ω_t
β	im Abschnitt 3.3: Geschwindigkeit des <i>mean-reverting</i> Prozesses und ein Parameter in den ARCH-Modellen im Abschnitt 4.3: positiver Parameter
β_t	Parameter der Prognose
δ	im Abschnitt 2.5: stetiger Mischungsparameter

	im Abschnitt 3.3: Parameter im ARMA-Prozess
δ_1	Parameter des zeitgleichen Volumens
δ_2	Parameter, der den Erklärungsgehalt des nicht prognostizierbaren Volumens misst
δ_λ	Persistenzparameter der Volatilität
$\{\varepsilon_t\}$	normalverteilter stochastischer Prozess
ε_t	im Abschnitt 2.3: Residuen im Abschnitt 4.2: <i>white noise</i>
ε_{t-1}	Fehlterm der Vorperiode im ARCH-Modell
ϕ	im Abschnitt 3.4: gemeinsame Komponente des Reservationspreises in den Abschnitten 4.2 und 4.4: Persistenzparameter
ϕ_0	Parameter im ARV-Modell
ϕ_1	Persistenzparameter
ϕ_2	Parameter im ARV-Modell
γ	im Abschnitt 3.2: durchschnittliches Niveau der logarithmierten Informationsrate im Abschnitt 3.4: separate Persistenzquelle im Abschnitt 4.2: Parameter im bivariaten SV-Modell
γ_0	im Abschnitt 3.3: Parameter im EGARCH-Modell im Abschnitt 4.4: Konstante
γ_1	Beta Faktor
γ_2	<i>risk-return tradeoff</i>
γ^2	Momentanvarianz des Prozesses $\ln \sigma^2(t)$
η	Störterm
η_t	im Abschnitt 3.3: standardnormalverteilte Zufallsvariable im Abschnitt 4.2: <i>white noise</i>
$\iota_{t,i,j}$	Indikatorvariable
$\tilde{\kappa}$	Parameter der Asymmetrie der Sensibilität
κ_t	Koeffizient, der proportional zum K_t ist
κ_τ	Exzess
λ	im Abschnitt 4.2: Parameter im bivariaten SV-Modell
λ_t	im Abschnitt 3.3: logarithmierte Varianz des ARV-Modells zum Zeitpunkt t

	im Abschnitt 4.3: logarithmierte Informationsrate zum Zeitpunkt t
λ	Vektor der latenten Variablen
μ	Erwartungswert
π_t	Trend in den Daten
θ	Mischungsvariable
θ_0	Parameter im EGARCH-Modell
θ_1	Parameter, der den Einfluss der verzögerten Renditeinnovationen auf die bedingte Varianz misst
θ_2	Parameter der Asymmetrie
θ	Parametervektor
ρ	im Abschnitt 3.3: Korrelationsparameter im Abschnitt 4.2: Parameter des <i>white noise</i> c_t
ρ_τ	Autokorrelationsfunktion
σ_t	Standardabweichung zum Zeitpunkt t
σ_t^2	Varianz zum Zeitpunkt t
σ_ϕ^2	Varianz des gemeinsamen Schocks
σ_ψ^2	Varianz des spezifischen Schocks
ζ	Parameter im bivariaten SV-Modell
τ	im Abschnitt 2.4: <i>lag</i> Länge im Abschnitt 4.2: Parameter im bivariaten SV-Modell
ν	Freiheitsgrade der Student- t -Verteilung
ω	im Abschnitt 2.4: Frequenz im Abschnitt 4.2: Parameter im bivariaten SV-Modell
$\{\omega_t\}$	ein latenter Prozess
ξ	Teilpreisänderung
ψ	spezifische Komponente des Reservationspreises
\mathfrak{I}_t	Informationsmenge zum Zeitpunkt t

Abkürzungen:

ACF	<i>autocorrelation function</i>
ARCH	<i>autoregressive conditional heteroscedasticity</i>
ARFIMA	<i>autoregressive fractionally integrated moving average</i>
ARMA	<i>autoregressive moving average</i>
ARV	<i>autoregressive volatility</i>
BHHH	Bernd-Hall-Hall-Hausman Algorithmus
Cov	Kovarianz
EGARCH	<i>exponential generalized autoregressive conditional heteroscedasticity</i>
FIGARCH	<i>fractionally integrated generalized autoregressive conditional heteroscedasticity</i>
GARCH	<i>generalized autoregressive conditional heteroscedasticity</i>
GMM	<i>generalized method of moments</i>
<i>i.i.d.</i>	<i>independently and identically distributed</i>
LM	Lagrange-Multiplikator
LR	Likelihood-Ratio-Statistik
MC	Monte-Carlo Methode
MDH	<i>mixed distribution hypothesis</i>
ML	Maximum Likelihood
MMS	<i>minimum mean square</i>
NIG	normal-inverse Gauß-Verteilung
QML	Quasi Maximum Likelihood
SARV	<i>stochastic autoregressive volatility</i>
SML	<i>simulated</i> Maximum Likelihood
SV	<i>stochastic volatility</i>
Var	Varianz
ZGWS	Zentraler Grenzwertsatz

1. Einführung

Zahlreiche Modelle in der Theorie der Finanzmärkte basieren auf einer Annahme über die Verteilungseigenschaften der Preisprozesse. Oft wird die Normalverteilung angenommen, weil sie durch den Erwartungswert und die Varianz vollständig beschrieben ist, was den Rechenvorgang erleichtert. Dabei kann die Normalverteilung die Finanzzeitreihen aufgrund deren besonderen Eigenschaften nur schlecht beschreiben. Die Abbildung 1 zeigt die empirische Verteilung der logarithmierten Renditen der Andritz AG Aktie über den Zeitraum von 2001 bis 2009.

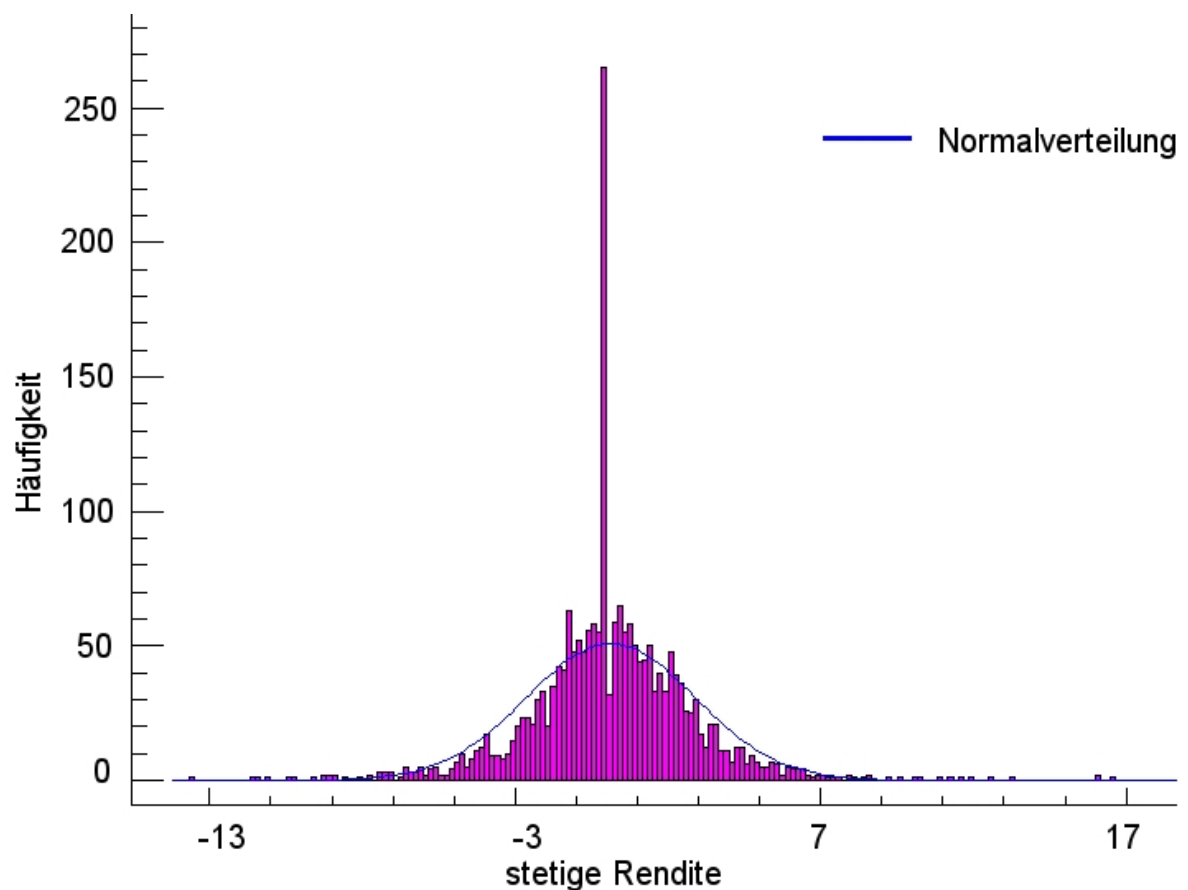


Abbildung 1: Logarithmierte Renditen von Andritz AG (2001-2009) im Vergleich zur Normalverteilung.
Quelle: Eigene Darstellung.

Im Vergleich mit der Normalverteilung wird ersichtlich, dass mehr Ausprägungen um den Erwartungswert und auf den *tails* der empirischen Verteilung liegen. Diese Eigenschaft ist

Leptokurtosis, sie ist beobachtbar in den empirischen Verteilungen der Renditen und Handelsvolumina.

Ein Ansatz zur Erklärung der Leptokurtosis in Renditeverteilungen liegt im Bereich von *behavioural finance*: Starke Preisbewegungen werden auf den ausgeprägten Optimismus bzw. Pessimismus der Marktteilnehmer zurückgeführt.

Der erste quantitative Ansatz stammt von Mandelbrot, der 1963 ein Lévy-stabiles nicht-Gauss Modell zur Modellierung der leptokurtischen Verteilung eines logarithmierten Preisprozess eingesetzt hat. Modelle der ARCH-Klasse stellen einen Ansatz zur Erklärung von *fat tails* in der Volatilität der Aktien dar.

Diese Arbeit konzentriert sich auf die Mischungsverteilungshypothese, die 1973 von Clark ausgearbeitet wurde und die einen weiteren Ansatz für die plausible Erklärung der leptokurtische Gestalt der Renditeverteilungen liefert. Im Modell von Clark werden Normalverteilungen mit gleichen Erwartungswerten und unterschiedlich großen Varianzen gemischt, wobei die Anzahl der am Tag eingetroffenen preisrelevanten Informationen als Mischungsvariable verwendet wird. Da die Anzahl der Informationen nicht direkt beobachtbar ist, wird eine Proxyvariable verwendet. Dafür schlägt Clark das Handelsvolumen vor und behandelt es als eine statische exogene Variable.

Viele Arbeiten (vgl. Andersen (1996), Mestel (2008) und Tauchen und Pitts (1983)) weisen darauf hin, dass das Handelsvolumen im Rahmen des Mischungsverteilungsmodells eine endogene Variable darstellt, das heißt es wird teilweise von den gleichen Faktoren wie die Varianz der Preisänderungen beeinflusst.

Diese Arbeit konzentriert sich auf die Entwicklung des Mischungsverteilungsmodells unter Berücksichtigung des endogenen Handelsvolumens und der dynamischen Informationsrate. Außerdem werden das *short-* und *long-run* Verhalten der Varianz der Preisänderungen und des Handelsvolumens und seine Implikationen für die Mischungsverteilungshypothese diskutiert.

Im zweiten Kapitel wird ein kurzer Überblick über die theoretischen Grundlagen gegeben.

Im dritten Kapitel werden das Standardmodell von Clark und seine Erweiterungen behandelt. Die Mischungsverteilungshypothese ist ein univariates Modell, in dem das Handelsvolumen als exogen betrachtet und die beobachtete Autokorrelation der Informationsrate nicht berücksichtigt wird. Der ARCH-Ansatz von Bollerslev et al. (1994) und der ARV-Ansatz von Taylor (1986) stellen eine dynamische Erweiterung des Mischungsverteilungsmodells von Clark dar. Einen statischen bivariaten Ansatz liefern Tauchen und Pitts (1983). Andersen (1996) berücksichtigt in seinem Modell die dynamischen Eigenschaften der Prozesse und betrachtet das Handelsvolumen als eine endogene Variable. Dieses Modell passt sich gut an die Daten an, liefert aber einen deutlich kleineren Persistenzparameter, als das univariate ARV-Modell.

Das vierte Kapitel befasst sich mit den möglichen Erklärungen für den Rückgang des Persistenzparameters in bivariaten Modellen. Liesenfeld (1998), Andersen (1996) und Mestel (2008) vermuten, dass die Informationsrate aus heterogenen Komponenten besteht. Liesenfeld (2001) führt neben der latenten Informationsrate die latente Informationssensibilität ein, welche die Varianz der Preisänderungen mitbeeinflusst. Mestel (2008) untersucht die statischen und dynamischen Effekte in Renditezeitreihen wie Handelsvolumenzeitreihen.

Im fünften Kapitel werden die zentralen Aspekte dieser Arbeit zusammengefasst.

2. Theoretische Grundlagen

2.1 Einführung

Finanzzeitreihen weisen besondere Eigenschaften auf, die deren Modellierung erschweren. Dazu zählen beispielsweise Abweichungen von der Normalverteilung der Renditen und serielle Abhängigkeiten in deren Varianz. In Folge werden jene Eigenschaften von Rendite- und Volumenzeitreihen diskutiert, die für diese Arbeit besonders wichtig sind. Es wird gezeigt, wie Verteilungen gemischt werden können. Außerdem wird Grundlegendes über *short-* und *long-memory* Verhalten in den Zeitreihen präsentiert, das für die Erläuterungen in den nächsten Abschnitten von Bedeutung ist.

2.2 Statistische Eigenschaften der Renditezeitreihen

Renditezeitreihen werden als stochastischer Prozess angesehen, der durch die Eigenschaften seiner Momentfunktionen beschrieben wird. Höhere Momente der Renditezeitreihen unterscheiden sich von jenen eines normalverteilten Prozesses und erschweren somit die Modellierung. In Folge wird auf die Schiefe und die Kurtosis sowie die Volatilität der Verteilungen der Renditezeitreihen eingegangen.

Empirische Untersuchungen zeigen, dass normale wie logarithmierte Renditen in der Regel einen höheren Kurtosiskoeffizient als 3 aufweisen, siehe dazu Mandelbrot (1963), Tsay (2002) und Schröder (2002). Das bedeutet, dass die Verteilung mehr Wahrscheinlichkeitsmasse um den Erwartungswert und in den *tails* hat, als das bei einer Normalverteilung der Fall ist.

Tsay (2002) beobachtet, dass Schiefe keine große Hürde für die Modellierung darstellt.

Finanzzeitreihen weisen keine linearen Abhängigkeiten auf, aber man kann Perioden mit großen oder kleinen Preisschwankungen beobachten. Das deutet auf eine nicht-lineare Abhängigkeit im zweiten Moment hin. Dieses Phänomen wurde erstmals von Mandelbrot (1963) und Fama (1965) beobachtet und heißt *volatility clustering*. Somit ist die Varianz der

Verteilung über die Zeit nicht konstant. Wenn die Varianz von den vorherigen Werten beeinflusst wird, spricht man von der bedingten Heteroskedastizität.

Schröder (2002) beschreibt den einfachsten Fall bedingter Heteroskedastizität mit folgendem Modell:

$$Y_t = \varepsilon_t \text{ mit } \varepsilon_t \sim N(0, \sigma_t^2) \text{ und } \sigma_t^2 = a + b \varepsilon_{t-1}^2. \quad (2.1)$$

Nach diesem Modell ist die Variable Y ein normalverteilter stochastischer Prozess mit Erwartungswert Null und zeitabhängiger Varianz. Die Varianz wird durch die Fehlerterme der Vorperiode bestimmt. Das ist ein so genanntes „*autoregressive conditional heteroskedasticity*“-Modell (ARCH-Modell).

2.3 Statistische Eigenschaften der Handelsvolumenzeitreihen

Wie die Renditezeitreihen, stellen Volumenzeitreihen einen stochastischen Prozess dar. Bei Mestel (2008) findet sich eine detaillierte Untersuchung der Verteilungseigenschaften des Volumenprozesses.

In den meisten Fällen wird die Hypothese der Normalverteilung verworfen. Logarithmierte tägliche Zeitreihen weisen Kurtosiswerte signifikant größer als 3 auf, was mehr Wahrscheinlichkeitsmasse an den Verteilungsenden bedeutet. In dieser Eigenschaft ähnelt das Volumen den Renditezeitreihen. Unterschiedlich ist, dass sich Volumenzeitreihen als asymmetrisch erweisen. Mestel (2008) beobachtet ein deutlich rechtsschiefes Verhalten für die untersuchten Aktien des US-Marktes. Das bedeutet, dass Handelstage mit sehr hohem Handelsvolumen häufiger auftreten, als jene mit sehr niedrigem Handelsvolumen.

Das bringt die Vermutung nahe, dass eine Verteilung, die Asymmetrie und Leptokurtosis zu berücksichtigen vermag, für die Modellierung der Handelsvolumenzeitreihen besser als die Normalverteilung geeignet wäre.

Mestel (2008) kommt zum Schluss, dass die normal-inverse Gauß-Verteilung zu den besten Ergebnissen führt. Die Nullhypothese der NIG-Verteilung kann für alle untersuchten Teilmärkte am seltensten verworfen werden.

Anders als der Renditeprozess erweist sich der Volumenprozess im ersten Moment als hoch persistent. Mestel (2008) beobachtet positive und statistisch signifikante Autokorrelationen für alle Verzögerungen. Die hyperbolische Abnahme der Autokorrelationsfunktionen lässt vermuten, dass es sehr langfristige Abhängigkeiten in den Daten gibt.

Zur Untersuchung der Abhängigkeiten im zweiten Moment nimmt Mestel (2008) an, dass das logarithmierte Handelsvolumen einem stationären ARMA(p, q)-Prozess folgt:

$$\Phi(B)V_t = c + \Theta(B)\varepsilon_t. \quad (2.2)$$

Dabei ist c eine Konstante und B ein Backshift-Operator: $B_{x_t}^i = x_{t-i}$.

Mithilfe des Lagrange-Multiplikator-Tests (LM) kann der Prozess, dem die quadrierten Residuen ε_t^2 folgen, untersucht werden, um das Vorliegen der ARCH-Effekte in den Volumenzeitreihen zu bestätigen bzw. zu widerlegen. Im Zuge der Untersuchung wurde die Nullhypothese der Homoskedastizität für alle Fälle eindeutig abgelehnt. Das bedeutet, dass der Effekt des *volatility clustering* in den Volumenzeitreihen, ähnlich wie bei Renditen, vorzufinden ist.

2.4 Short memory und long memory Prozesse

Die Untersuchung von Mestel (2008) zeigt, dass sich die Abhängigkeiten in den Volumenzeitreihen für über *lag* 100 erstrecken. Diese langfristigen Abhängigkeiten weisen darauf hin, dass der Volumenprozess *long-memory* Eigenschaften aufweist.

Ob ein Prozess *short-memory* oder *long-memory* Eigenschaften besitzt kann mittels Analyse der Autokorrelationsfunktion (ACF) eines Prozesses überprüft werden. Im Falle eines stationären und invertierbaren ARMA-Prozesses nimmt die ACF exponentiell ab und die spektrale Dichte ist für alle Frequenzen begrenzt. Jene Prozesse weisen *long-memory*

Eigenschaften auf, deren ACF mit zunehmender Verzögerung langsam, hyperbolisch abschwächen und deren spektrale Dichte im Niederfrequenzbereich nicht begrenzt ist.

Eine formale Darstellung findet sich bei Pong et al. (2008) und Tsay (2002): Ein diskreter stochastischer Prozess $\{y_t\}$ mit der ACF ρ_τ und spektraler Dichte $f(\omega)$ besitzt *short-memory* Eigenschaften, wenn $\sum_{\tau=1}^n \rho_\tau$ konvergiert mit $n \rightarrow \infty$ und $f(\omega)$ für alle Frequenzen begrenzt ist; andernfalls wird ein *long-memory* Prozess vermutet.

Abbildung 2 zeigt das hyperbolische Abschwächen einer Autokorrelationsfunktion eines *long-memory* Prozesses am Beispiel des logarithmierten Handelsvolumens der Siemens-Aktie. Die Werte der empirischen ACF sind hoch signifikant auch bei *lag* 100.

Ein formales Beispiel eines long memory Prozesses ist das ARFIMA(p,d,q)-Modell (*autoregressive fractionally integrated moving average*). Sie stellt eine Verallgemeinerung des ARIMA-Modells dar:

$$\Phi(B)(1-B)^d X_t = c + \Theta(B)\varepsilon_t, \quad \varepsilon_t \sim i.i.d.(0, \sigma^2). \quad (2.3)$$

Anders als im einfachen ARIMA-Modell kann hier der Differenzparameter d auch nicht ganzzahlige Werte annehmen und dadurch die langfristigen Abhängigkeiten zum Ausdruck bringen. Die kurzfristigen Eigenschaften werden laut Mestel (2008) durch die ARMA-Polynome $\Phi(\cdot)$ und $\Theta(\cdot)$ abgebildet..

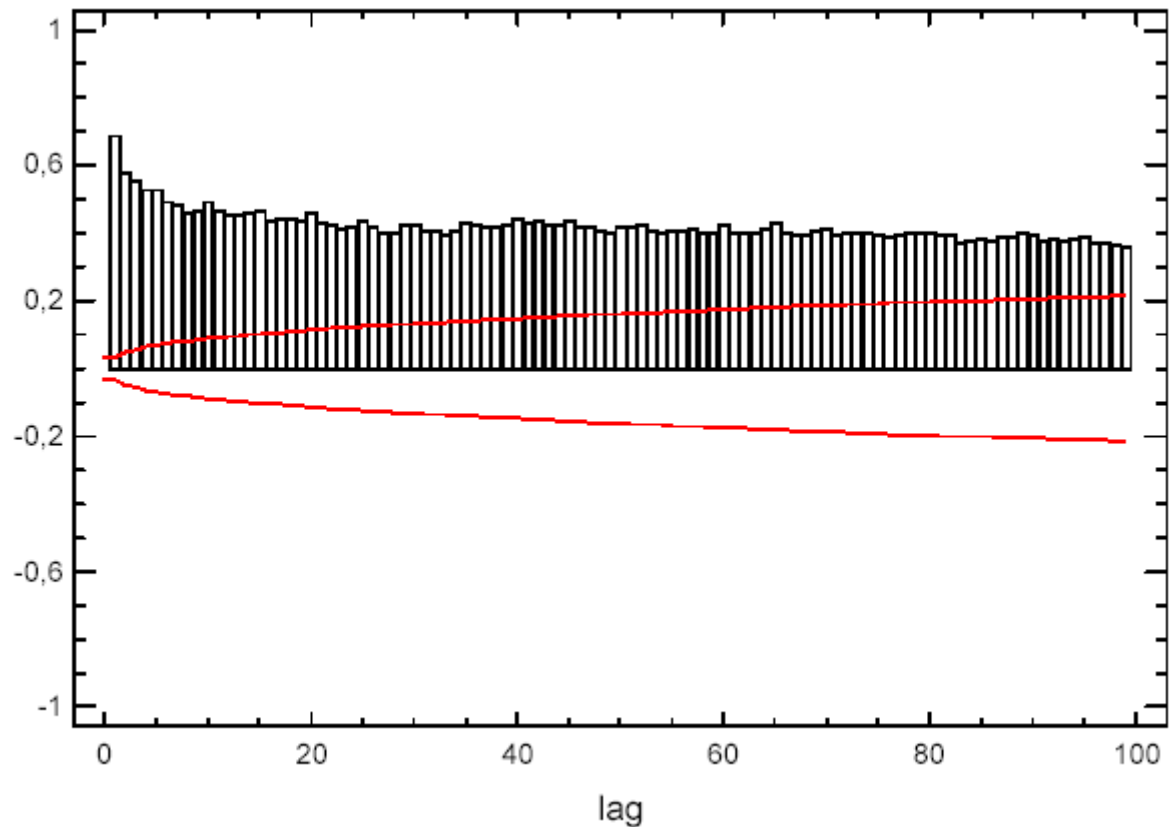


Abbildung 2: Empirische Autokorrelationsfunktion des logarithmierten Handelsvolumens der Siemens-Aktie, 1994 - 2005.

Quelle: Mestel (2008), Seite 43.

Mestel (2008) merkt an, dass Werte von $d \in (0;1)$ besonders interessant sind: In diesem Fall ist die ACF für alle Verzögerungen positiv und zeigt in Abhängigkeit von d ein hyperbolisches Abschwingen, jedoch sind Varianzfunktion und Kovarianzfunktion nur für $d \in (0;0,5)$ stationär. Werden für $d \in (0;0,5)$ Autokorrelationen ρ_τ und spektrale Dichten $f(\omega)$ durch entsprechend τ^{2d-1} und ω^{-2d} dividiert, konvergieren die Ergebnisse gegen positive Konstanten bei $\tau \rightarrow \infty$ und $\omega \rightarrow \infty$. Tabelle 1 zeigt Eigenschaften der Autokorrelationsfunktionen und der spektralen Dichten für *short-memory* sowie *long-memory* Prozesse.

	<i>Short memory</i>	<i>Long memory</i>
Autokorrelation $\tau \rightarrow \infty$	$\frac{ \rho_\tau }{\phi^\tau} \rightarrow C_1 > 0,$ $0 < \phi < 1$	$\frac{\rho_\tau}{\tau^{2d-1}} \rightarrow D_1 > 0,$ $d > 0$
Frequenz $\omega \rightarrow \infty$	$f(\omega) \rightarrow C_2 > 0$	$\frac{f(\omega)}{\omega^{-2d}} \rightarrow D_2 > 0,$ $d > 0$

Tabelle 1: Eigenschaften der Autokorrelationsfunktionen und der spektralen Dichten.

Quelle: Pong et al. (2008), Seite 28.

Dabei stellen C_1 , C_2 , D_1 und D_2 positive Konstanten dar.

2.5 Mischungsverteilungen

Da sich diese Arbeit mit der Mischungsverteilungshypothese befasst, wird in diesem letzten Abschnitt der theoretischen Grundlagen das Konzept der Mischungsverteilungen präsentiert.

Bei der „Mischung“ von Verteilungen wird über Verteilungen der gleichen Klasse gemittelt. Ein diskretes Beispiel einer Mischung der Normalverteilungen mit Erwartungswerten μ_1 und μ_2 sowie Varianzen σ_1^2 und σ_2^2 findet sich bei Ronning (2003):

$$g(x|\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \alpha) = (1 - \alpha)f_N(x|\mu_1, \sigma_1^2) + \alpha f_N(x|\mu_2, \sigma_2^2), \quad (2.4)$$

dabei ist $g(\cdot)$ die Mischungsverteilung, $f_N(\cdot)$ die Dichtefunktionen der Normalverteilungen und $0 \leq \alpha \leq 1$. Ist σ_1^2 klein und σ_2^2 vergleichsweise groß, dann sind bei $\alpha = 0,05$ 95% der Preisänderungen gemäß $N(\mu_1, \sigma_1^2)$ und die restlichen 5% gemäß $N(\mu_2, \sigma_2^2)$ verteilt. Dank der hohen σ_2^2 kann mehr Wahrscheinlichkeitsmasse auf den *tails* der Verteilung konzentriert werden. Der kleine Prozentsatz der Preisänderungen, die mit $N(\mu_2, \sigma_2^2)$ verteilt sind, bedeutet, dass die meisten Preisänderungen einfach normalverteilt sind. Dabei ist die Mischungsverteilung keine Normalverteilung mehr. Sie kann auch zweigipfelig sein und eine beliebig große endliche Varianz haben.

Abbildung 3 zeigt ein diskretes Beispiel einer Mischungsverteilung, deren Gleichung folgendermaßen aussieht:

$$g(x|0, 6, 1, 6, 0,25) = 0,75 f_N(x|0,0) + 0,25 f_N(x|6,6).$$

Die Mischungsverteilung, die sich daraus ergibt, ist zweigipfelig und hat mehr Wahrscheinlichkeitsmasse auf den *tails*, als eine Normalverteilung.

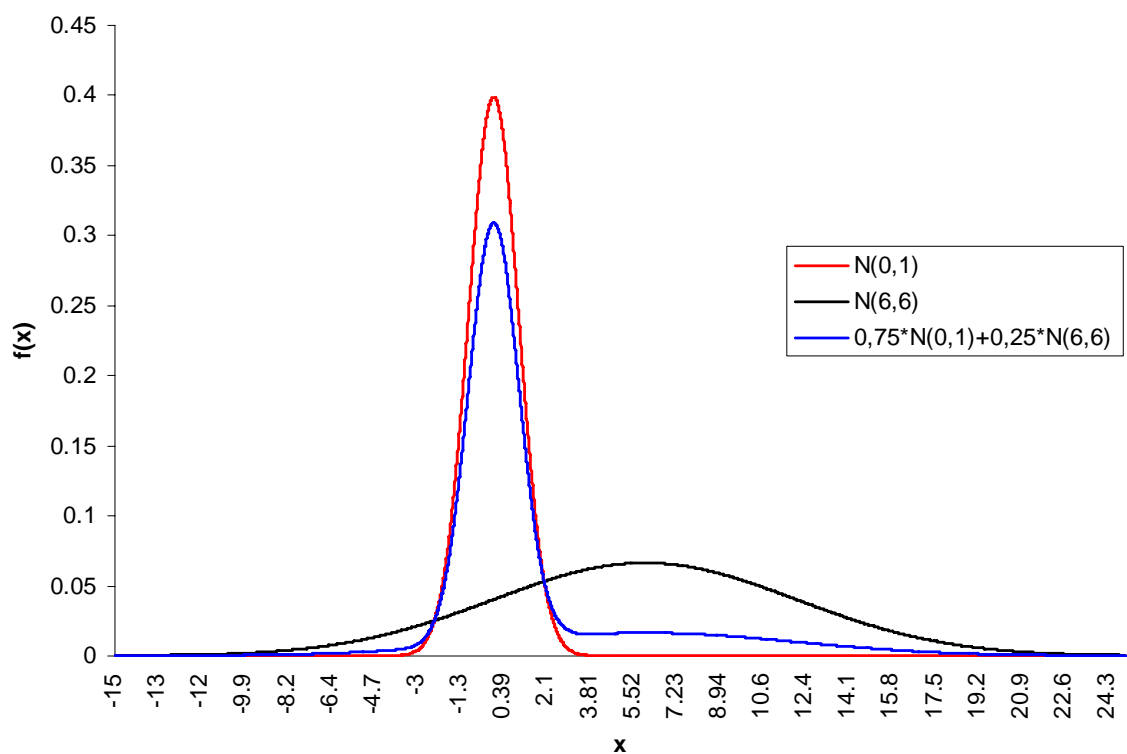


Abbildung 3: Diskretes Beispiel einer Mischungsverteilung.

Quelle: Eigene Darstellung.

Im Falle einer stetigen Mischung wird über alle möglichen Erwartungswerte unter Vorgabe einer Varianz gemittelt. Ein Beispiel der stetigen Mischungsverteilung findet sich bei Ronning (2003):

$$g(x|\gamma, \delta) = \int f(x|\theta, \gamma) m(\theta|\delta) d\theta. \quad (2.5)$$

Die Dichtefunktion $f(x|\theta, \gamma)$ hängt von Parametern θ und γ ab. Dabei ist $m(\theta|\delta)$ die Dichte der Mischungsverteilung mit der Mischungsvariable θ und Mischungsparameter δ . Das bedeutet, dass es über den Parameter θ für den vorgegebenen Parameter δ gemittelt wird.

Der Vorteil der Mischungsverteilungen besteht laut Tsay (2002) darin, dass sie wie Normalverteilungen interpretiert werden können, endliche hohe Momente besitzen und die Leptokurtosis nachbilden können. Nachteilig ist, dass sich die Schätzung der Mischungsparameter als schwierig erweist.

2.6 Zusammenfassung

In diesem Kapitel wurde ein kurzer Überblick über die theoretischen Grundlagen gegeben. Es wurde die statische und dynamische Heteroskedastizität, die für Renditezeitreihen und Handelsvolumenzeitreihen typisch ist, diskutiert. Außerdem wurde das Konzept der Mischungsverteilung erläutert und auf die Eigenschaft von *short* und *long memory* in stochastischen Prozessen eingegangen.

Im nächsten Kapitel werden diese Konzepte zur Erläuterung der Mischungsverteilungshypothese und deren Erweiterungen eingesetzt.

3. Entwicklung der Mischungsverteilungshypothese

3.1 Einführung

Im zweiten Kapitel wurde gezeigt, dass empirische Renditeverteilungen Eigenschaften aufweisen, die auf den ersten Blick mit theoretischen Einschätzungen nicht übereinstimmen: Die Preisänderungen einer Periode können als Summe der Teilpreisänderungen dargestellt werden, und nach der Hypothese der effizienten Märkte sind diese Teilpreisänderungen von einander unabhängig. Weisen die Daten eine gleiche Verteilung auf und sind sie von einander unabhängig, kommt der Zentrale Grenzwertsatz zur Anwendung¹. Gemäß dem ZGWS konvergiert die Summe von *i.i.d.* Verteilungen gegen eine Standardnormalverteilung. Wie im zweiten Kapitel dieser Arbeit gezeigt wurde, weichen aber die empirischen Renditeverteilungen signifikant von einer Normalverteilung zugunsten einer leptokurtischen Verteilung ab.

Die Mischungsverteilungshypothese stellt eine empirisch relevante Erklärung für die leptokurtische Gestalt der empirischen Verteilung dar.

Alle Varianten des Standardmodells können in statische und dynamische sowie in univariate und bivariate Modelle unterteilt werden.

	Univariat	Bivariat
Statisch	Clark (1973)	Tauchen und Pitts (1983)
Dynamisch	Bollerslev et al. (1994)	Andersen (1996)

Tabelle 2: Das Standardmodell und seine Erweiterungen.

Quelle: Eigene Darstellung.

In diesem Kapitel wird zunächst das Modell von Clark sowie seine Erweiterungen präsentiert.

¹ Die Anforderungen an den Erwartungswert und die Varianz sollen erfüllt sein.

3.2 Das Standardmodell von Clark

Die Intensität des Handels ist nicht an allen Handelstagen gleich. Die Änderungen der Geschwindigkeit der Preisentwicklung führt Clark in seinem 1973 erstellten Modell darauf zurück, dass nicht immer gleich viel preisrelevante Information auf den Markt eintrifft.

Er lässt den logarithmierten Kurs p_t folgendermaßen darstellen²:

$$p_t = p_{l_t}^*, t \in \mathbb{N}, l \in \mathbb{N} \quad (3.1)$$

Dabei gibt der Index t die Kalenderzeit in Tagen und der Index l die Transaktionszeit an, beide sind hier diskret. Also entspricht $p_{l_t}^*$ dem Preisprozess gemessen auf der Ereignisskala.

Dabei ist der Prozess l_t der steuernde Prozess und p_t ist dem Prozess $p_{l_t}^*$ untergeordnet. Wird die tägliche Preisänderung mit r_t bezeichnet, ist die Verteilung $r_t(p_t)$ der Verteilung $r_t(p_{l_t}^*)$ untergeordnet.

Die Preisänderung einer Periode ist die Summe der Teilpreisänderungen ξ . Die tägliche Preisänderung r_t lässt sich wie folgt darstellen:

$$r_t = \sum_{s=l_{t-1}+1}^{l_t} \xi_s \quad \text{mit} \quad \xi_s = p_s^* - p_{s-1}^*, \quad (3.2)$$

beziehungsweise für die Anzahl der Informationen $n_t = l_t - l_{t-1}$

$$r_t = \sum_{i=1}^{n_t} \xi_{l_{t-1}+i} = \sum_{i=1}^{n_t} \xi_{t,i} \quad (3.3)$$

Die Anzahl der preisrelevanten Informationen variiert von Tag zu Tag und bestimmt den täglichen Zuwachs im Prozess n_t . Die täglichen Preisänderungen ergeben sich laut Liesenfeld (1998) aus der Summe der Teilpreisänderungen $\xi_{t,i}$, wobei die Anzahl der Summanden n_t stochastisch ist.

Da die Anzahl der Summanden n_t eine stochastische Größe darstellt, kann der herkömmliche ZGWS nicht verwendet werden. Um die Verteilung der Preisänderungen approximieren zu

² Aus Konformitätsgründen wird die Notation aus Liesenfeld (1998) übernommen.

können, verallgemeinert Clark (1973) den ZGWS mit zwei Annahmen hinsichtlich der Teilpreisänderungen und der täglichen Anzahl der Informationen. Seine erste Annahme ist, dass die Variablen $\xi_{t,i}$ und n_t stochastisch unabhängig sind. Des Weiteren nimmt Clark (1973) an, dass die Teilpreisänderungen *i.i.d.* mit Erwartungswert Null und Varianz σ_ξ^2 sind.

Liesenfeld (1998) zerlegt die Anzahl an Informationen in zwei Komponenten:

$$n_t = \lceil \bar{n} w_t \rceil. \quad (3.4)$$

Dabei stellt \bar{n} eine positive Konstante dar und w_t ist eine von $\xi_{t,i}$ unabhängige positive Zufallsvariable.

Laut Clark (1973) kann die bedingte Verteilung für tägliche Preisänderungen gegeben die Anzahl der Informationen n_t für hinreichend großen \bar{n} durch die folgende Normalverteilung approximiert werden:

$$r_t | n_t \sim N(0, \sigma_\xi^2 n_t). \quad (3.5)$$

Somit ist die Varianz dieser Normalverteilung proportional zur stochastischen Anzahl der Informationen n_t . Liesenfeld (1998) stellt die täglichen Preisänderungen in der Regressionsschreibweise folgendermaßen dar:

$$r_t = \sigma_\xi \sqrt{n_t} u_t \quad \text{mit} \quad u_t = \frac{1}{\sigma_\xi \sqrt{n_t}} \sum_{i=1}^{n_t} \xi_{t,i} \sim i.i.d.N(0,1). \quad (3.6)$$

Dabei bezeichnet u_t die Innovationen.

Es wird angenommen, dass $f(r_t | n_t)$ die Dichtefunktion der in (3.5) angegebenen bedingten Normalverteilung ist und $g(n_t)$ die Randdichte der Anzahl der Informationen. Dann ergibt sich die Dichte der unbedingten Verteilung $f(r_t)$ aus der Summe bzw. dem Integral der mit $g(n_t)$ gewichteten bedingten Verteilungen $f(r_t | n_t)$ (vgl. Abschnitt 2.5). Für einen diskreten Fall gilt

$$f(r_t) = \sum_{n_t} f(r_t | n_t) g(n_t). \quad (3.7)$$

Wird n_t als eine stetige Zufallsvariable modelliert, ergibt sich entsprechend

$$f(r_t) = \int f(r_t | n_t)g(n_t)dn_t . \quad (3.8)$$

Die unbedingte Verteilung der täglichen Preisänderungen ist somit eine Mischung der bedingten Normalverteilungen mit unterschiedlichen Varianzen. Als Mischungsvariable dient die Anzahl der auf den Markt eintreffenden Informationen n_t .

Aus der Gleichung (3.6) leitet Liesenfeld (1998) die unbedingten Momente der unbedingten Verteilung $f(r_t)$ ab:

$$E(r_t) = \sigma_\xi E(\sqrt{n_t}u_t) = \sigma_\xi E(\sqrt{n_t})E(u_t) = 0 \quad (3.9)$$

$$\text{Var}(r_t) = \sigma_\xi^2 E(n_t u_t^2) = \sigma_\xi^2 E(n_t)E(u_t^2) = \sigma_\xi^2 E(n_t) \quad (3.10)$$

$$E(r_t^3) = \sigma_\xi^3 E(n_t^{3/2} u_t^3) = \sigma_\xi^3 E(n_t^{3/2})E(u_t^3) = 0 \quad (3.11)$$

$$E(r_t^4) = \sigma_\xi^4 E(n_t^2 u_t^4) = \sigma_\xi^4 E(n_t^2)E(u_t^4) = 3\sigma_\xi^4 E(n_t^2). \quad (3.12)$$

Clark (1973) beweist, dass die Kurtosis von $f(r_t)$ eine steigende Funktion der Varianz von n_t ist, wenn p_t^* normalverteilt und *i.i.d.* ist, und n_t *i.i.d.* ist und eine endliche Varianz besitzt, die unabhängig von p ist.

Die Wölbung μ^4 / σ^4 einer Normalverteilung beträgt 3, deswegen ergibt sich der bedingte Erwartungswert als

$$E(r_t^4 | n_t) = 3n_t^2 \sigma_\xi^4. \quad (3.13)$$

Der unbedingte Erwartungswert beträgt

$$E_{n_t} (3n_t^2 \sigma_\xi^4) = 3\sigma_\xi^4 (E(n_t)^2 + \text{Var}(n_t)). \quad (3.14)$$

Daraus resultiert die Gleichung für den Exzess κ_τ :

$$\kappa_\tau = \frac{3\sigma_\xi^4 (E(n_t)^2 + \text{Var}(n_t))}{E(n_t)^2 \sigma_\xi^4} - 3 = 3 \frac{E(n_t)^2 + \text{Var}(n_t)}{E(n_t)^2} - 3 = 3 \frac{\text{Var}(n_t)}{E(n_t)^2}. \quad (3.15)$$

Analysiert man die unbedingte Verteilung, so wird aus der Gleichung (3.11) ersichtlich, dass die Schiefe der unbedingten Verteilung $f(r_t)$ gleich Null ist, das heißt, die Verteilung ist symmetrisch. Der Exzess κ_τ in der Gleichung (3.15) ist positiv, das bedeutet, dass die Verteilung leptokurtische Gestalt aufweist. Dabei ist der Exzess direkt proportional zur Varianz der Anzahl der Informationen n_t pro Tag. Diese Verteilungseigenschaft kann dadurch erklärt werden, dass sich die Mischungsverteilung aus einem Mittel ergibt. Das Mittel wird aus einer Normalverteilung mit großer Varianz, die an Tagen mit verhältnismäßig hoher Informationsintensität zu erwarten ist, und aus einer Normalverteilung mit geringerer Varianz gebildet.

Das Mischungsverteilungsmodell von Clark (1973) liefert eine plausible theoretische Erklärung für die empirische Leptokurtosis der Finanzzeitreihen. Da die Anzahl der am Tag eingetroffenen preisrelevanten Informationen nicht direkt beobachtbar ist, bedient er sich einer Proxyvariable für die empirische Überprüfung des Modells. Dafür wird das tägliche Handelsvolumen herangezogen. Es wird angenommen, dass das tägliche Handelsvolumen proportional zur Anzahl der Informationen ist. Im Rahmen des Mischungsverteilungsmodells soll die Varianz der Preisänderungen positiv mit dem Handelsvolumen korrelieren.

Zur Überprüfung wird der Zusammenhang zwischen den Kursänderungen und dem Handelsvolumen $v(t)$ geschätzt:

$$r_t^2 = h(v_t) + \varepsilon_t. \quad (3.16)$$

Dabei ist r_t^2 das Maß für die Varianz des Tages t und die Funktion $h(v_t)$ stellt eine Approximation für die latente Informationsrate dar. Clark (1973) hat zwei Spezifikationen dieser Form untersucht:

$$h(v_t) = \alpha v_t^\beta \quad (3.17)$$

und

$$h(v_t) = a e^{bv_t}. \quad (3.18)$$

Für die beiden Spezifikationen wird die Kurtosis geschätzt und ein Test auf Normalverteilung bzw. Lognormalverteilung durchgeführt. Clark (1973) kommt zum Schluss, dass mit der Spezifikation (3.17) die beste Anpassung an die Rohdaten erreicht wird. Es zeigt sich, dass das Handelsvolumen einen signifikant positiven Einfluss auf die Varianz der Preisänderungen hat, was zugunsten der Mischungsverteilungshypothese spricht. Die Ergebnisse zeigen außerdem, dass v_t der untersuchten Stichprobe lognormal und nicht normal verteilt ist.

Wenn das Handelsvolumen als Proxyvariable für die Anzahl der Informationen verwendet wird, wird es als eine exogene Variable behandelt. In der Literatur findet man zahlreiche empirische und theoretische Untersuchungen, die zum Schluss führen, dass der Prozess des Handelsvolumens (zumindest teilweise) von gleichen Faktoren wie der Prozess der Varianz bestimmt wird³. Somit handelt es sich jedoch um eine endogene Variable. Betrachtet man das Handelsvolumen als exogen, werden die Annahmen des Standardmodells verletzt. Außerdem ist die Schätzung des Zusammenhangs zwischen den Kursänderungen und dem Handelsvolumen mit Simultanitätsbias behaftet. Somit ist die Verwendung des Handelsvolumens als eine Proxyvariable im Standardmodell nicht sinnvoll.

Alternativ kann zur empirischen Überprüfung des Modells eine Annahme über die Verteilung der Anzahl der Informationen n_t unterstellt werden. Dabei wären nur solche Verteilungen sinnvoll, die negativen Werten von n_t keine positive Wahrscheinlichkeit zuordnen würden. Laut Tauchen und Pitts (1983) erweist sich die Lognormalverteilung in empirischen Untersuchungen als eine besser geeignete, im Vergleich zur Poissonverteilung. Clark (1973)

³ Siehe dazu z.B. Andersen (1996) und Mestel (2008).

schlägt vor, die Anzahl an Informationen als eine identisch und unabhängig lognormalverteilte Zufallsvariable zu modellieren. Dann ergibt sich laut Liesenfeld (1998) folgende empirische Spezifikation für das Mischungsverteilungsmodell:

$$r_t = \mu + \sigma_\varepsilon \sqrt{n_t} u_t \quad (3.19)$$

mit $u_t \sim i.i.d.N(0,1)$ und $\ln(n_t) \sim i.i.d.N(\gamma, v^2)$.

Dabei fungiert v als eine Mischungsvariable und γ gibt das durchschnittliche Niveau von $\ln(n_t)$ an. Zu beachten ist, dass die Preisänderungen r_t hier als seriell unabhängig modelliert werden, da unterstellt wird, dass n_t und u_t seriell unabhängig sind.

Die Parameter des so spezifizierten Modells können mit der Maximum-Likelihood-Methode (ML), sowie mit der Verallgemeinerten-Momenten-Methode (*generalized method of moments*, GMM) geschätzt werden.

Liesenfeld (1998) verwendet folgende Spezifikation der Log-Likelihoodfunktion für das Standardmodell (3.19):

$$\ln L(\mathbf{r}; \boldsymbol{\theta}) = \sum_{t=1}^T \ln f(r_t; \boldsymbol{\theta}). \quad (3.20)$$

Dabei ist $f(r_t; \boldsymbol{\theta})$ die Dichte der unbedingten Verteilung von r_t und $\boldsymbol{\theta} = (\mu, \gamma, v)'$ der zu schätzende Parametervektor.

Laut Becker (1998) wird mit der GMM versucht, einen Vektor, der abhängige beobachtbare Variablen beinhaltet, möglichst gut durch eine Matrix der unabhängigen Variablen zu erklären. Der Schätzer für den Parametervektor $\boldsymbol{\theta}$ wird so gewählt, dass die Differenz zwischen den beobachteten Momenten der abhängigen Variablen und den aus dem zu schätzenden Modell erwarteten Momenten minimiert wird.

Liesenfeld (1998) spezifiziert den GMM-Schätzer wie folgt:

$$\hat{\theta}_{\text{GMM}} = \arg \min_{\theta} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{m}(r_t, \theta) \right]' \mathbf{W} \left[\frac{1}{T} \sum_{t=1}^T \mathbf{m}(r_t, \theta) \right]. \quad (3.21)$$

Dabei ist $\mathbf{m}(r_t, \theta)$ das Stichprobenmittel und \mathbf{W} eine $(q \times q)$ -dimensionale Gewichtungsmatrix.

Liesenfeld (1998) erhält ein System aus $p = 3$ zu schätzenden Parametern und $q = 5$ sich aus den Momenten $E(r_t)$, $E(|r_t - \mu|)$, $\text{Var}(r_t)$, $E(|r_t - \mu|^3)$ und $E(|r_t - \mu|^4)$ ergebenden Restriktionen:

$$\frac{1}{T} \sum_{t=1}^T \mathbf{m}(r_t, \theta) = \frac{1}{T} \sum_{t=1}^T \begin{pmatrix} r_t - \mu \\ |r_t - \mu| - \sqrt{\frac{2}{\pi}} \exp\left\{\frac{1}{2}\gamma + \frac{1}{8}v^2\right\} \\ (r_t - \mu)^2 - \exp\left\{\gamma + \frac{1}{2}v^2\right\} \\ |r_t - \mu|^3 - 2\sqrt{\frac{2}{\pi}} \exp\left\{\frac{3}{2}\gamma + \frac{9}{8}v^2\right\} \\ (r_t - \mu)^4 - 3\exp\{2\gamma + 2v^2\} \end{pmatrix}. \quad (3.22)$$

Liesenfeld (1998) schlägt das iterative Verfahren für die GMM-Schätzung vor. Dabei wird als erste konsistente aber nicht effiziente Schätzung eine Einheitsmatrix als Gewichtungsmatrix verwendet. Sie wird zur Ermittlung der optimalen Gewichtungsmatrix eingesetzt. Diese wird im zweiten Schritt zur Ermittlung der effizienten GMM-Schätzung herangezogen. Das Verfahren wird solange wiederholt, bis die Differenz zwischen zwei aufeinander folgenden Schätzungen hinreichend gering ist.

Aus den empirischen Ergebnissen der Untersuchungen von Liesenfeld (1998) ist ersichtlich, dass die theoretische Dichtefunktion, die sich aus der ML-Schätzung ergibt, die Leptokurtosis der empirischen Verteilung wesentlich besser beschreibt, als eine Normalverteilung. Somit kann das Mischungsverteilungsmodell zumindest einen Teil der Leptokurtosis der Preisänderungen erklären. Das Ergebnis für die Parameterschätzung mit der Verallgemeinerten-Momenten-Methode ist, dass die Hypothese einer richtigen Spezifikation des Mischungsverteilungsmodells zu keinem der üblichen Signifikanzniveaus abgelehnt werden kann. Dabei ist es unerheblich, ob die möglichen Autokorrelationen in den

Momentrestriktionen bei der Schätzung berücksichtigt werden, – die Aussage über die Angemessenheit des Mischungsverteilungsmodells bleibt erhalten.

Der Vergleich der geschätzten Parameter zeigt, dass die Ergebnisse aus den ML- und GMM-Schätzverfahren nahe bei einander liegen, obwohl beide Methoden auf unterschiedlichen a-priori Informationen basieren.

3.3 Dynamische Erweiterung des univariaten Modells

Das Standardmodell von Clark (1973) liefert eine plausible Erklärung für die leptokurtische Gestalt der Verteilung der Preisänderungen. Das oft anzutreffende Phänomen des Volatilität-Clusterings, also das autoregressive Verhalten der Varianz, kann das Modell nicht erklären, weil für die Anzahl der Informationen und somit für die Varianz keine Autokorrelation zugelassen wird.

Grundsätzlich ist das beobachtete Verhalten der Varianz von Preisänderungen mit dem Mischungsverteilungsmodell vereinbar. Die Preisänderungen ergeben sich im Standardmodell aus (vgl. Gleichung 3.16)

$$r_t = \mu + \sigma_\xi \sqrt{n_t} u_t \quad \text{mit} \quad u_t \sim i.i.d.N(0,1), \quad (3.23)$$

dabei sind die Informationsrate n_t und die Innovationen u_t seriell unabhängig. Durch e_t werden die Residuen der Preisänderungen bezeichnet:

$$e_t = r_t - \mu = \sigma_\xi \sqrt{n_t} u_t \quad (3.24)$$

und

$$e_t^2 = (r_t - \mu)^2 \quad (3.25)$$

wird als lokales Maß für die Varianz des Tages t eingesetzt. Liesenfeld (1998) zeigt, dass eine positive Autokorrelation der Informationsrate in einer positiven Autokorrelation der quadrierten Residuen resultiert, wenn die Autokorrelationsfunktion der quadrierten Residuen der Preisänderungen aus dem Mischungsverteilungsmodell gebildet wird:

$$\begin{aligned} \text{Cov}(e_t^2, e_{t-\tau}^2) &= E\left[\left(\sigma_\xi^2 n_t u_t^2 - E(\sigma_\xi^2 n_t u_t^2)\right) \times \left(\sigma_\xi^2 n_{t-\tau} u_{t-\tau}^2 - E(\sigma_\xi^2 n_{t-\tau} u_{t-\tau}^2)\right)\right] \\ &= \sigma_\xi^4 \text{Cov}(n_t, n_{t-\tau}) \end{aligned} \quad (3.26)$$

mit $\tau = 1, 2, \dots$.

Somit ist das Mischungsverteilungsmodell, das die serielle Korrelation der Informationsrate berücksichtigt, mit dem beobachteten Effekt von *volatility clustering* grundsätzlich vereinbar. Da diese serielle Abhängigkeit in der Varianz der Preisänderungen beobachtet wird, sollte Clarks Annahme der Unabhängigkeit abgelehnt werden.

Die Plausibilität der positiven Autokorrelation in der Informationsrate kann durch mehrere empirische und theoretische Ansätze erklärt werden. Eine mögliche theoretische Erklärung für *volatility clustering* ist, dass die preisrelevanten Informationen in Clustern auf den Markt kommen. Eine andere Erklärung ist, dass die Investoren die Informationen nicht gleichzeitig, sondern sequentiell erhalten. Durch die Änderung der Nachfrage ergeben sich die Preisänderungen. Die Preise kommen zum Gleichgewicht, wenn alle Investoren die neuen Informationen erhalten und verarbeitet haben. Noch eine Erklärung liefern die mikroökonomischen Modelle, die auf der Informationsasymmetrie basieren. Der Handel der besser informierten Marktteilnehmer führt zu einer verzögerten Aufdeckung der privaten Informationen und somit zur Persistenz im Preisänderungsprozess.

Die so begründete Autokorrelation in der Informationsrate und somit in der Varianz der Preisänderungen kann mit Hilfe von ARV (*autoregressive volatility*) und ARCH (*autoregressive conditional heteroskedasticity*) Modellen nachgebildet werden.

3.3.1 Das ARV-Modell

Liesenfeld (1998) untersucht das von Taylor 1986 vorgeschlagene ARV-Modell. Es stellt eine zeitdiskrete Approximation des zeitstetigen Itô-Prozesses dar, der in Modellen zur Bewertung

von Optionen eine wichtige Rolle spielt. Die geometrische Brownsche Bewegung ist ein Diffusionsprozess, der am häufigsten für die zeitstetige Modellierung eines stetigen Wertpapierkurses $s(t)$ verwendet wird:

$$ds(t) = \alpha s(t)dt + \sigma s(t)dw(t), \quad t_0 \leq t \leq T. \quad (3.27)$$

Dabei ist

α erwartete momentane Rendite des Wertpapiers;

σ^2 Momentanvarianz;

$w(t)$ standardisierte Brownsche Bewegung.

Für den logarithmierten Wertpapierkurs $p(t) = \ln s(t)$ erhält Liesenfeld (1998) nach der Anwendung des Satzes von Itô:

$$dp(t) = \left(\alpha - \frac{1}{2} \sigma^2 \right) dt + \sigma dw(t). \quad (3.28)$$

Aus dieser Gleichung ergibt sich eine Normalverteilung für die Preisänderungen, berechnet aus der Veränderung des Prozesses $p(t)$ über ein festes Zeitintervall. Außerdem ist die Varianz im Zeitablauf konstant. Diese Spezifikation ist mit den empirischen Gegebenheiten – Leptokurtosis und im Zeitablauf variierende Varianz – nicht konsistent. Aus diesem Grund betrachtet Liesenfeld (1998) die Momentanvarianz σ^2 als eine stochastische Variable und modelliert sie als Prozess $\ln \sigma^2(t)$. Liesenfeld (1998) benutzt folgende Gleichungen zur stetigen Modellierung des Preisprozesses und der Momentanvarianz unter Verwendung von einer zweidimensionalen Brownschen Bewegung ($w_1(t), w_2(t)$) mit Korrelationsparameter ρ :

$$dp(t) = \left[\alpha - \frac{1}{2} \sigma^2(t) \right] dt + \sigma(t)dw_1(t) \quad (3.29)$$

$$d \ln \sigma^2(t) = \beta [\bar{\alpha} - \ln \sigma^2(t)] dt + \gamma dw_2(t) \quad (3.30)$$

$$dw_1(t)dw_2(t) = \rho dt. \quad (3.31)$$

Dabei ist

- γ^2 Momentanvarianz vom Prozess $\ln \sigma^2(t)$;
 $\bar{\alpha}$ langfristiger Mittelwert vom Prozess $\ln \sigma^2(t)$;
 β Geschwindigkeit des *mean-reverting* Prozesses.

Die Variablen sind nur in diskreter Zeit beobachtbar, daher müssen die Gleichungen (3.29) – (3.31) approximiert werden. Liesenfeld (1998) schlägt die einfachste Form der diskreten Approximation, das Euler-Verfahren vor. Dabei wird eine Schrittweite Δt der Diskretisierung gewählt und der Beobachtungszeitraum $[t_0, T]$ in N Intervalle zerlegt, sodass $\Delta t = (T - t_0)/N$. Nach dieser Approximation ergibt sich ein System aus folgenden zeitdiskreten Prozessen:

$$p_t = p_{t-1} + \left(\alpha - \frac{1}{2} \sigma_{t-1}^2 \right) \Delta t + \sigma_{t-1} \sqrt{\Delta t} u_t, \quad (3.32)$$

$$\ln \sigma_t^2 = \ln \sigma_{t-1}^2 + \beta (\bar{\alpha} - \ln \sigma_{t-1}^2) \Delta t + \gamma \sqrt{\Delta t} \eta_t. \quad (3.33)$$

Dabei sind u_t und η_t standardnormalverteilte Zufallsvariablen, die keine Korrelation aufweisen und laut Gleichung (3.31) eine Kovarianz von ρ besitzen. Die logarithmierte Varianz σ_t^2 folgt gemäß Gleichung (3.33) einem AR(1)-Prozess, und ihre verzögerten Werte beeinflussen sowohl die Preisänderungen $r_t = p_t - p_{t-1}$ als auch die Varianz selber.

Nelson (1990; zit. n. Liesenfeld 1998, S. 45) hat gezeigt, dass die diskreten Gleichungen (3.32) und (3.33) für immer kleiner werdende Schrittweite der Diskretisierung Δt gegen die stetigen Prozesse (3.29) und (3.30) konvergieren.

Liesenfeld (1998) setzt $\Delta t = 1$ und ersetzt den verzögerten Wert σ_{t-1} in der Gleichung (3.32) durch σ_t . Des Weiteren vernachlässigt er die Abhängigkeit des Erwartungswertes $\left(\alpha - \frac{1}{2} \sigma_{t-1}^2 \right) \Delta t$ der Preisänderungen von der verzögerten Varianz σ_{t-1}^2 und erhält aus den Gleichungen (3.32) und (3.33) die folgende reparametrisierte Spezifikation:

$$r_t = \mu + \sigma_t u_t \quad \text{mit } u_t \sim i.i.d.N(0,1), \quad (3.34)$$

$$\ln \sigma_t^2 = \phi_0 + \phi_1 \ln \sigma_{t-1}^2 + \phi_2 \eta_t \quad \text{mit } \eta_t \sim i.i.d.N(0,1). \quad (3.35)$$

Die Gleichungen (3.34) und (3.35) stellen das ARV-Modell dar, das auch als SV (*stochastic volatility*) Modell bezeichnet wird.

Die Gleichung (3.32) konvergiert gegen den Diffusionsprozess (3.29) im Sinne der schwachen Konvergenz sowohl für σ_{t-1} als auch für σ_t . Das heißt, dass beide Spezifikationen als diskrete Approximationen des zeitstetigen Preisprozesses angesehen werden können. Laut Gleichung (3.35) folgen die logarithmierten Varianzen einem AR(1)-Prozess. Da der Störterm η_t sowie die Innovation u_t unbeobachtbar sind, wird σ_t als eine unbeobachtbare Zufallsvariable modelliert. Die einzige Variable, die beobachtbar ist, ist die Preisänderung r_t .

Die Spezifikation der Preisänderungen r_t im ARV-Modell (3.34) ist mit jener des Standardmodells von Clark konsistent (3.19), wenn der Korrelationsparameter ρ in der Gleichung (3.31) gleich Null ist. Der nicht prognostizierbare Teil der Preisänderungen besteht jeweils aus dem Produkt zweier Zufallsvariablen, $\sigma_t u_t$ im ARV-Modell und $\sigma_\xi \sqrt{n_t} u_t$ im Standardmodell respektive. In beiden Modellen sind die Innovationen $u_t \sim i.i.d.N(0,1)$, also entspricht die Standardabweichung σ_t des ARV-Modells dem Term $\sigma_\xi \sqrt{n_t}$ des Mischungsverteilungsmodells von Clark. Eine von Clarks Annahmen war, dass $\sigma_\xi \sqrt{n_t}$ und u_t stochastisch unabhängig sind. Für den Fall $\rho = 0$ im ARV-Modell sind der Störterm η_t und die Innovation u_t , und somit u_t und σ_t , laut Liesenfeld (1998) stochastisch unabhängig.

Für $\sigma_t^2 = \sigma_\xi^2 n_t$ erhält Liesenfeld (1998) folgende Spezifikation für die Anzahl der täglichen Informationen:

$$\ln(\sigma_\xi^2 n_t) = \phi_0 + \phi_1 \ln(\sigma_\xi^2 n_{t-1}) + \phi_2 \eta_t, \quad (3.36)$$

also folgt die logarithmierte Anzahl der Informationen n_t laut dem ARV-Modell einem AR(1)-Prozess mit normalverteiltem Störterm η_t . Im Falle, dass der Faktor ϕ_1 gleich Null ist und dementsprechend keine Autokorrelation von n_t modelliert wird, ist die Zufallsvariable $\ln(n_t)$ *i.i.d.* normalverteilt. Die Gleichung (3.36) repräsentiert also eine Verallgemeinerung der von Clark (1973) vorgeschlagenen Spezifikation für die Informationsrate, vgl. Gleichung (3.19).

Die Preisänderungen, die sich aus dem ARV-Modell (3.34) und (3.35) ergeben, lassen sich folgendermaßen darstellen:

$$r_t | \sigma_t^2 \sim N(\mu, \sigma_t^2), \quad (3.37)$$

$$\ln \sigma_t^2 | \ln \sigma_{t-1}^2 \sim N([\phi_0 + \phi_1 \ln \sigma_{t-1}^2], \phi_2^2). \quad (3.38)$$

Für den Fall, dass der Parameter ϕ_1 betragsmäßig kleiner Eins ist, sind die logarithmierten Varianzen $\ln \sigma_t^2$ kovarianz-stationär mit einer unbedingten Verteilung, die laut Liesenfeld (1998) mit

$$\ln \sigma_t^2 \sim N(\bar{y}, \bar{v}^2) = N\left(\frac{\phi_0}{1-\phi_1}, \frac{\phi_2^2}{1-\phi_1^2}\right) \quad (3.39)$$

gegeben ist. Es ist offensichtlich, dass die logarithmierten Varianzen σ_t^2 einer unbedingten Normalverteilung folgen. Liesenfeld (1998) unterstreicht, dass die unbedingte Verteilung der Preisänderungen $f(r_t)$, die sich aus der Gleichung (3.37) ergibt, eine Mischungsverteilung aus Normalverteilungen mit einer lognormalverteilten Varianz als Mischungsvariable darstellt, wie im Standardmodell von Clark (1973). Das bedeutet, dass die unbedingten Momente von r_t im ARV-Modell die gleiche Form wie jene im Standardmodell aufweisen.

Der Exzess ergibt sich als

$$\kappa_\tau = 3 \frac{\text{Var}(\sigma_t^2)}{\text{E}(\sigma_t^2)^2} = 3(\exp\{\bar{v}^2\} - 1), \quad (3.40)$$

das heißt, dass die unbedingte Verteilung der Preisänderungen $f(r_t)$ für $\bar{v} > 0$ leptokurtisch ist.

Liesenfeld (1998) hält für das ARV-Modell Folgendes fest: Einerseits sind die Preisänderungen bzw. die Residuen $e_t = r_t - \mu$ unkorreliert, es gibt aber positive Autokorrelation der quadrierten Residuen e_t^2 bei $\phi_1 > 0$. Formal dargestellt

$$\rho_e(\tau) = 0 \quad \text{für } \tau = 1, 2, \dots \quad (3.41)$$

$$\rho_{e^2}(\tau) = \frac{\exp\{\bar{v}^2 \phi_1^\tau\} - 1}{3 \exp\{\bar{v}^2\} - 1} \quad \text{für } \tau = 1, 2, \dots \quad (3.42)$$

dabei sind $\rho_e(\tau)$ und $\rho_{e^2}(\tau)$ die Autokorrelationsfunktionen der Residuen und der quadrierten Residuen respektive, und τ ist die *lag* Länge. Dabei ist $\rho_{e^2}(\tau)$ für eine gegebene τ umso größer, je höher der Parameter ϕ_1 ist. Außerdem nimmt $\rho_{e^2}(\tau)$ mit zunehmender *lag* Länge für $|\phi_1| < 1$ ab. Liesenfeld (1998) kommt zum Schluss, dass sich der Parameter ϕ_1 als Maß für die Stärke der seriellen Korrelation, also für die Persistenz von Schocks im Varianzprozess eignet.

Zusammenfassend lässt sich festhalten, dass das ARV-Modell eine Verallgemeinerung des Standardmodells darstellt und mit den empirischen Eigenschaften der Preisänderungen, also der Leptokurtosis sowie den seriellen Korrelationen in quadrierten Residuen, vereinbar ist.

3.3.2 ARCH-Modelle

In ARCH-Modellen wird die bedingte Varianz als eine Funktion der verzögerten Variablen abgebildet. Der Preisprozess kann mithilfe eines ARCH-Modells folgendermaßen dargestellt werden:

$$r_t = \mu + \sqrt{h_t} z_t \quad (3.43)$$

mit $z_t \sim i.i.d. D(0,1)$ und $h_t = Var(r_t | I_{t-1})$.

Dabei ist

μ der prognostizierbare Teil der Preisänderungen;

z_t Innovationen, hier eine *i.i.d.* Zufallsvariable mit $E(z_t) = 0$ und $Var(z_t) = 1$;

h_t bedingte Varianz von r_t zum Zeitpunkt t gegeben die Informationsmenge I_{t-1} ;

D eine noch nicht spezifizierte Verteilung, Annahme über welche in weiterer Folge getroffen wird.

Diese Informationsmenge enthält die ganze beobachtbare Information zum Zeitpunkt $t-1$. Liesenfeld (1998) nimmt an, dass diese Information aus der Menge der vergangenen Preisänderungen besteht, also $I_{t-1} = \{r_1, \dots, r_{t-1}\}$. Somit wird die bedingte Varianz h_t als eine Funktion der verzögerten Preisänderungen r_1, \dots, r_{t-1} dargestellt.

Um das ARCH-Modell zu spezifizieren, muss die Wahl der Spezifikation für die bedingte Varianz h_t und eine Annahme über die Verteilung der Innovationen z_t getroffen werden.

ARCH(q)

In der Literatur gibt es mehrere Spezifikationen für die bedingte Varianz h_t . Im linearen ARCH(q)-Modell von Engle (1982) ist die bedingte Varianz eine Funktion der vergangenen quadrierten Innovationen:

$$h_t = \alpha_0 + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \dots + \alpha_q z_{t-q}^2. \quad (3.44)$$

Bollerslev et al. (1994) schreiben die Gleichung (3.44) für weitere Umformung folgendermaßen um:

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i z_{t-i}^2 \equiv \alpha_0 + \alpha(B)z_{t-1}^2. \quad (3.45)$$

Dabei müssen die Parameter α_0 und α_i Folgendes erfüllen, damit die bedingte Varianz positiv ist: $\alpha_0 > 0$; $\alpha_1 \geq 0, \dots, \alpha_q \geq 0$. B bezeichnet den Backshift-Operator.

Bollerslev et al. (1994) definieren $v_t \equiv z_t^2 - h_t$ und schreiben das ARCH(q)-Modell aus der Gleichung (3.45) wie folgt um:

$$z_t^2 = \alpha_0 + \alpha(B)z_{t-1}^2 + v_t. \quad (3.46)$$

Da $E_{t-1}(v_t) = 0$, entspricht das Modell einem AR(q)-Modell für quadrierte Innovationen z_t^2 . Die Innovationen sind zwar nicht autokorreliert, aber auch nicht unabhängig über die Zeit. Wegen der quadrierten Innovationen kann dieses Modell das für die Finanzzeitreihen typische Verhalten nachbilden, indem extreme Ausprägungen ungeachtet des Vorzeichens aufeinander folgen.

GARCH(p,q)

Die empirische Anwendung von ARCH-Modellen verlangt oft nach einer höheren Parametrisierung und nach der Modellierung größerer *lags* für die Spezifizierung der bedingten Varianz. Das verallgemeinerte ARCH(p,q)-Modell (*generalized* ARCH, GARCH) von Bollerslev (1986; zit. n. Bollerslev et al. 1994, S. 2968) kommt diesen Anforderungen nach:

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i z_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \equiv \alpha_0 + \alpha(B)z_{t-1}^2 + \beta(B)h_{t-1}. \quad (3.47)$$

Um eine positive bedingte Varianz aus diesem GARCH(p,q)-Modell zu erhalten, müssen alle Koeffizienten des entsprechenden linearen ARCH Modells positiv sein. Für ein einfaches GARCH(1,1)-Modell ist laut Bollerslev et al. (1994) eine positive bedingte Varianz bei Parameterbeschränkungen $\alpha_0 \geq 0$, $\alpha \geq 0$ und $\beta \geq 0$ fast immer gewährleistet.

Das Modell (3.47) spezifiziert die bedingte Varianz als eine lineare Funktion der q verzögerten quadrierten Residuen und p vergangenen quadrierten bedingten Varianzen. Dieses Modell ist einem ARCH-Modell ähnlich, bietet aber mehr Flexibilität bei der Parametrisierung der bedingten Varianz. Bollerslev et al. (1994) merken an, dass das einfache GARCH(1,1)-Modell ausreichend ist, um die Autokorrelation in der Varianz anzupassen. Mit Hilfe dieses Modells lassen sich die für Finanzzeitreihen typischen Eigenschaften des *fat tails* und des *volatility clustering* gut nachbilden. Das Modell kann auch für die Berücksichtigung

solcher Unregelmäßigkeiten wie erwartete Veröffentlichung der Informationen oder auch handelsfreie Tage angepasst werden. Da die Residuen in Gleichung (3.47) quadriert in die bedingte Varianz einfließen, hängt diese nur von deren Größe und nicht von deren Vorzeichen ab. Somit kann das GARCH-Modell den Leverage-Effekt nicht nachbilden.

Bollerslev et al. (1994) beschreiben den Leverage-Effekt als die Tendenz zur negativen Korrelation zwischen den gegenwärtigen Preisänderungen und der zukünftigen Volatilität. Dieses Phänomen ist teilweise auf fixe Kosten und finanzielles und operationelles Leverage zurückzuführen. Normalerweise steigt das Leverage, wenn der Wert des Eigenkapitals eines verschuldeten Unternehmens fällt. Das erhöht die Volatilität der Eigenkapitalrendite bei der gleich bleibenden Rendite des Unternehmens.

EGARCH(p,q)

Der Leverage-Effekt findet Berücksichtigung in exponentiellen GARCH(p,q)-Modell (EGARCH) von Nelson (1991). Laut diesem Modell folgt die logarithmierte bedingte Varianz $\{\ln(h_t)\}$ einem ARMA(p,q)-Prozess⁴:

$$\ln h_t = \delta_0 + \sum_{i=1}^q \delta_i \ln h_{t-i} + \sum_{j=1}^p \gamma_j g(z_{t-j}). \quad (3.48)$$

Um den Leverage-Effekt nachzubilden, muss $g(z_t)$ eine Funktion des Wertes sowie des Vorzeichens von z_t sein. Nelson (1991) schlägt eine lineare Kombination aus z_t und $|z_t|$ vor:

$$g(z_t) \equiv \theta_0 z_t + \gamma_0 [|z_t| - E(|z_t|)]. \quad (3.49)$$

Die Funktion $g(z_t)$ besteht also aus zwei Komponenten, $\theta_0 z_t$ und $\gamma_0 [|z_t| - E(|z_t|)]$, beide haben einen Erwartungswert von Null. Falls z_t eine symmetrische Verteilung besitzt, sind diese Komponenten orthogonal, obwohl nicht unabhängig. Die Funktion $g(z_t)$ ist linear in z_t auf dem Abschnitt $0 < z_t < \infty$ mit der Steigung $\theta + \gamma$, und auf dem Abschnitt $-\infty < z_t \leq 0$ mit der Steigung $\theta - \gamma$. Somit wird eine asymmetrische Reaktion des Varianzprozesses auf die Preissteigerung bzw. Preissenkung gewährleistet.

⁴ Im Interesse der Kontinuität wird hier die Notation von Bollerslev et al. (1990) beibehalten.

In diesem Modell sind keine Parameterbeschränkungen notwendig, um eine positive bedingte Varianz zu gewährleisten. Die einfache Berücksichtigung von Residuen über $g(\cdot)$ lässt den Einfluss deren Vorzeichen auf die logarithmierte Varianz zu. Für $\gamma_j \theta_0 < 0$ bedeutet das, dass sinkende Preise zu einer höheren Varianz führen, während bei steigenden Preisen eine niedrigere Varianz zu erwarten ist.

Bollerslev et al. (1994) merken an, dass eine zeitabhängige Modellierung des Parameters δ_0 eine Berücksichtigung der vorhersehbaren Ereignisse oder der handelsfreien Perioden ermöglichen würde.

ARCH- t

Nach der Spezifizierung der bedingten Varianz h_t muss eine Annahme über die Verteilung der Innovationen z_t getroffen werden. Die aus den empirischen Daten berechneten Innovationen \hat{z}_t weisen eine Verteilung mit leptokurtischer Gestalt auf. Somit ist das ARCH-Modell mit einer Normalverteilungsannahme für die Innovationen (ARCH-N) für die Modellierung ungeeignet. Liesenfeld (1998) modelliert die Innovationen mithilfe einer standardisierten t -Verteilung (ARCH- t), welche die Leptokurtosis besser beschreiben kann. Eine ARCH- t -Verteilung ist eine Mischung von Normalverteilungen mit variierenden Varianzen $z_t | \sigma_{z_t}^2 \sim N(0, \sigma_{z_t}^2)$. Die Varianzen $\sigma_{z_t}^2$ dienen als Mischungsvariable und folgen einer invertierten Gammaverteilung. Liesenfeld (1998) normiert die unbedingte Varianz von z_t gemäß (3.13) auf 1 und erhält folgende Dichtefunktion für die unbedingte Verteilung von z_t :

$$f(z_t) = [\pi(w-2)]^{-1/2} \frac{\Gamma\left(\frac{w+1}{2}\right)}{\Gamma\left(\frac{w}{2}\right)} \left[1 + \frac{z_t^2}{w-2}\right]^{-(w+1)/2}, \quad \text{mit } w > 2. \quad (3.50)$$

Dabei ist w die Anzahl der Freiheitsgrade und $\Gamma(\cdot)$ die Gammafunktion. Für immer kleiner werdende $1/w$ konvergiert diese standardisierte t -Verteilung gegen eine $N(0,1)$ -Verteilung.

ARCH-N-LN

Hsieh (1989) hat eine empirische Untersuchung für fünf Fremdwährungen durchgeführt, wobei er versucht hat, solche Spezifikation eines ARCH-Modells zu finden, welche die beste Anpassung an die Daten ermöglicht. Er ist zum Schluss gekommen, dass mit nicht-normalen Verteilungen bessere Ergebnisse erzielt werden, darunter auch mit der Normal-Lognormal-Mischung, also einer Mischung aus Normalverteilungen mit einer lognormalverteilten Mischungsvariable.

Liesenfeld (1998) spezifiziert das ARCH-N-LN-Modell wie folgt:

$$z_t | \sigma_{z_t}^2 \sim N(0, \sigma_{z_t}^2), \quad \text{mit} \quad \ln(\sigma_{z_t}^2) \sim N(b_0, b_1^2). \quad (3.51)$$

Die Dichtefunktion der unbedingten Verteilung von z_t ergibt sich in folgender Form:

$$f(z_t) = \int_{-\infty}^{+\infty} (2\pi \exp\{b_0 + b_1 \varepsilon_t\})^{-1/2} \exp\left\{-\frac{z_t^2}{2 \exp\{b_0 + b_1 \varepsilon_t\}}\right\} \times (2\pi)^{-1/2} \exp\left\{-\frac{\varepsilon_t^2}{2}\right\} d\varepsilon_t. \quad (3.52)$$

Setzt man den Parameter $b_0 = 0$, entspricht die Gleichung (3.52) der Dichte einer Normalverteilung.

Vergleicht man die im ARCH-Modell spezifizierte bedingte Varianz h_t wiederum mit dem Term $\sigma_{\xi}^2 n_t$ des Standardmodells von Clark, wird ersichtlich, dass sie nur unter unrealistischen Annahmen gleichzusetzen sind. Liesenfeld (1998) merkt an, dass h_t im ARCH-Modell eine Funktion der Preisänderungen r_1, \dots, r_{t-1} modelliert wird, das heißt, dass die bedingte Varianz schon vor Beginn des Handelstages t realisiert ist. Im Gegensatz dazu ist $\sigma_{\xi}^2 n_t$ eine Variable, die erst am Ende des Handelstages realisiert wird, da eine preisrelevante Information jederzeit während des Handelstages auftreten kann.

Liesenfeld (1998) leitet aus dem ARCH-Modell folgende Spezifikation für die bedingte Varianz ab:

$$h_t = \text{Var}(r_t | I_{t-1}). \quad (3.53)$$

Das bedeutet, dass die bedingte Varianz h_t eine Prognose für die Anzahl der preisrelevanten Informationen darstellt, basierend auf den vergangenen Preisänderungen I_{t-1} .

Des Weiteren zeigt Liesenfeld (1998), dass die Innovationen gemäß dem Standardmodell von Clark unter realistischen Annahmen nicht normalverteilt sind. Er stellt die Innovationen wie folgt dar:

$$z_t = \frac{r_t - \mu}{\sqrt{h_t}} \quad (3.54)$$

und leitet folgende Gleichung für die Innovationen im Standardmodell ab:

$$z_t = \frac{r_t - \mu}{\sqrt{\text{Var}(r_t | I_{t-1})}} = \sqrt{\frac{n_t}{\text{E}(n_t | I_{t-1})}} u_t. \quad (3.55)$$

Im Mischungsverteilungsmodell von Clark ist u_t normalverteilt. Wenn n_t nicht exakt durch I_{t-1} prognostiziert wird, können die Innovationen z_t nicht normalverteilt sein. Gemäß Gleichung (3.55) folgen die Innovationen z_t vielmehr einer leptokurtischen Verteilung, die sich laut Liesenfeld (1998) aus einer Mischung von Normalverteilungen ergibt:

$$z_t | m_t \sim \text{N}(0, m_t) \quad \text{mit} \quad m_t = \frac{n_t}{\text{E}(n_t | I_{t-1})}. \quad (3.56)$$

Taylor (1994; zit. n. Liesenfeld 1998, S. 53) schlägt vor, die Gleichung der Preisänderungen im ARCH-Modell folgendermaßen zu formulieren:

$$r_t = \mu + \sqrt{h_t} z_t = \mu + \sqrt{h_t} \sqrt{m_t} u_t \quad (3.57)$$

mit $u_t \sim i.i.d.\text{N}(0,1)$ und $\text{E}(m_t) = 1$,

so dass

$$z_t | m_t \sim N(0, m_t) \quad \text{und} \quad z_t \sim D(0,1). \quad (3.58)$$

In Verbindung mit Liesenfelds (1998) Interpretation der bedingten Varianz als eine Prognose der Anzahl der preisrelevanten Informationen (3.53) kann m_t in den Gleichungen (3.57) und (3.58) als Fehler dieser Prognose gesehen werden. Liesenfeld (1998) merkt an, dass bei $m_t > 1$ die Anzahl der preisrelevanten Informationen n_t durch h_t überschätzt wird und umgekehrt. Im ARCH- t -Modell folgt der Prognosefehler m_t einer invertierten Gammaverteilung, im ARCH-N-LN-Modell einer Lognormalverteilung. Im ARCH-N-Modell, das wegen seiner Normalverteilungsannahme für die Modellierung ungeeignet ist, würde m_t einer degenerierten Ein-Punkt-Verteilung folgen, sodass $m_t = 1$ für alle t ist. Das würde bedeuten, dass die Anzahl der Informationen für jeden Zeitpunkt durch die bedingte Varianz richtig geschätzt worden wäre.

Unter Berücksichtigung der Tatsache, dass u_t in den Gleichungen (3.19) und (3.57) eine *i.i.d.* $N(0,1)$ verteilte Variable ist, leitet Liesenfeld (1998) folgende Entsprechung ab:

$$\sigma_\xi^2 n_t = h_t m_t. \quad (3.59)$$

Um die Auswirkungen der ARCH-Modelle auf das Verhalten der Informationsrate n_t zu analysieren, setzt Liesenfeld (1998) entsprechende Spezifikationen in die Gleichung (3.59) ein. Für das GARCH-Modell erhält er

$$\sigma_\xi^2 n_t = \left[\alpha_0 + \alpha_1 (h_{t-1} m_{t-1} u_{t-1}^2) + \beta_1 h_{t-1} \right] m_t = \alpha_0 m_t + \frac{m_t}{m_{t-1}} (\alpha_1 z_{t-1}^2 + \beta_1) \sigma_\xi^2 n_{t-1}. \quad (3.60)$$

Für das EGARCH-Modell (1,1) gilt Folgendes (vgl. Gleichung (3.48)):

$$\begin{aligned} \ln(\sigma_\xi^2 n_t) &= \delta_0 + \delta_1 \ln h_{t-1} + \gamma_1 g(z_{t-1}) + \ln m_t \\ &= \delta_0 + \delta_1 \ln(\sigma_\xi^2 n_{t-1}) + \gamma_1 g(z_{t-1}) + \ln m_t - \delta_1 \ln m_{t-1}. \end{aligned} \quad (3.61)$$

Aus den Gleichungen (3.60) und (3.61) wird ersichtlich, dass die Anzahl der preisrelevanten Informationen n_t in den ARCH-Modellen vom vergangenen Wert n_{t-1} abhängt, das heißt, dass die serielle Abhängigkeit in der Informationsrate modelliert wird. Deswegen stellen ARCH-Modelle eine dynamische Erweiterung des statischen Standardmodells dar. Des weiteren hängt n_t laut den Gleichungen (3.60) und (3.61) von den Prognosefehlern m_t und m_{t-1} ab.

Zusammenfassend kann man festhalten, dass die bedingte Varianz h_t in den ARCH-Modellen als eine Funktion der Informationsmenge $I_{t-1} = \{r_1, \dots, r_{t-1}\}$ modelliert wird. Die Informationsrate n_t hängt von eigenen verzögerten Werten ab. Somit stellt das ARCH-Modell, genau wie das ARV-Modell, eine dynamische Erweiterung des statischen Standardmodells dar. Im Gegensatz zum ARV-Modell hängt die Informationsrate im ARCH-Modell auch von den verzögerten Preisänderungen ab.

Der Nachteil des ARCH-Modells besteht laut Liesenfeld (1998) darin, dass die exogene Informationsrate als eine latente Größe modelliert wird, was eine fundierte empirische Überprüfung nicht zulässt. Dabei wird aus dem Verhalten der Preisänderungen auf die Verteilungseigenschaften der Informationsrate geschlossen. Für eine Überprüfung soll in einem univariaten Modell eine geeignete Proxyvariable wie das Handelsvolumen miteinbezogen werden; in einem bivariaten Modell soll das gemeinsame Verhalten der Preisänderungen und anderer Variablen (wie wiederum das Handelsvolumen) untersucht werden, um Rückschlüsse auf das Verhalten der Informationsrate ziehen zu können.

3.3.3 Schätzung der Modelle und Vergleich der empirischen Ergebnisse

Liesenfeld (1998) bedient sich der ML-Methode zur Schätzung der Modelle der ARCH-Klasse. Aufgrund der Autokorrelation im nicht beobachtbaren Varianzprozess kann die gleiche Methode für die Schätzung des ARV-Modells nicht verwendet werden. Deswegen greift Liesenfeld (1998) auf die SML(*simulated maximum likelihood*)-Methode zurück, die einen anschließenden Modellvergleich ermöglicht.

Für eine empirische Überprüfung schätzt Liesenfeld (1998) das GARCH(1,1)- und das EGARCH(1,1)-Modell und vergleicht die Ergebnisse bei der Verwendung einer

Normalverteilung, einer Student- t -Verteilung und einer Lognormal-Normal-Mischung als bedingte Verteilung.

Für die Modelle der ARCH-Klasse kann eine Likelihoodfunktion vergleichsweise einfach ermittelt werden. Sie ergibt sich auch einem Produkt der bedingten Dichtefunktionen der Preisänderungen. Eine Log-Likelihoodfunktion hat das Maximum an der gleichen Stelle wie eine Likelihoodfunktion, ist aber einfacher zu berechnen. Liesenfeld (1998) ermittelt die Log-Likelihoodfunktion der ARCH-Modelle wie folgt:

$$\ln L(\mathbf{r}; \boldsymbol{\theta}) = \sum_{t=1}^T -\frac{1}{2} \ln h_t + \ln f\left(\frac{r_t - \mu}{h_t^{1/2}}\right), \quad (3.62)$$

dabei stellt \mathbf{r} einen Vektor der Realisationen der Preisänderungen $\mathbf{r} = \{r_t\}_{t=1}^T$ dar,

$\boldsymbol{\theta}$ ist der zu schätzende Parametervektor,

die bedingten Dichtefunktionen sind definiert durch die bedingten Verteilungen $r_t | I_{t-1} \sim D(\mu, h_t)$, und

$f(z_t)$ ist die Dichtefunktion des für die Innovationen $z_t = (r_t - \mu)/h_t^{1/2}$ unterstellten Verteilungsgesetzes.

Liesenfeld (1998) maximiert die Log-Likelihoodfunktion mit dem Berndt-Hall-Hall-Hausman-Algorithmus (BHHH) und ermittelt die geschätzten Standardfehler der Parameterschätzungen aus der Inversen der Hesse-Matrix. BHHH ist ein Schritt-Algorithmus, der zu nicht-linearen Schätzmethoden gehört. In der Hesse-Matrix werden die partiellen zweiten Ableitungen einer mehrdimensionalen Funktion zusammen gefasst.

Die Anpassungsgüte der EGARCH- und ARCH-Modelle überprüft Liesenfeld (1998) mithilfe des Kullback-Leibler-Kriteriums sowie des Akaike- und Schwarz-Kriteriums. Bei dem Kullback-Leibler-Kriterium werden die Log-Likelihoodfunktionen an der Stelle der Maximum-Likelihood-Schätzung verglichen. Dabei wird jenes Modell bevorzugt, das den höchsten Wert der Log-Likelihoodfunktion aufweist. Ein potentielles Problem dieses Informationskriteriums ist die Überparametrisierung, das heißt, tendenziell wird ein Modell mit vielen Parametern bevorzugt. Dieses Problem wird vom Akaike- und Schwarz-Kriterium umgangen indem die Anzahl der Modellparameter in die Entscheidung einfließt.

Die Persistenz der Varianzschocks hängt im GARCH-Modell von der Summe $\alpha_1 + \beta_1$ und im EGARCH-Modell vom Parameter δ_1 ab. Liesenfelds (1998) empirische Untersuchungen der ausgewählten DAX-Werte (Siemens, Daimler-Benz, Volkswagen, Deutsche Bank) zeigen, dass die Summe $\alpha_1 + \beta_1$ sowie der Parameter δ_1 für alle vier Aktien nah an den Wert 1 heranreichen, was eine hohe Persistenz im Varianzprozess der Aktien bedeutet.

Liesenfeld (1998) stellt fest, dass die empirisch geschätzte Anzahl der Freiheitsgrade für das GARCH- t -Modell sowie für das EGARCH- t -Modell zu niedrig ist, um die Student- t -Verteilungen durch eine Normalverteilung approximieren zu können. Ähnlich wie bei der Student- t -Verteilung, stellt eine Normal-Lognormal-Mischung eine signifikante Verbesserung gegenüber der Annahme der Normalverteilung dar.

Des Weiteren vergleicht Liesenfeld (1998) die Werte des Kullback-Leibler-Kriteriums sowie des Akaike- und Schwarz-Kriteriums für das GARCH- und EGARCH-Modell, jeweils unter der Annahme der Normalverteilung, Student- t -Verteilung und Normal-Lognormal-Verteilung. Der Vergleich zeigt, dass das EGARCH-Modell für gegebene Verteilungsannahmen einen höheren Wert der Log-Likelihoodfunktion aufweist, als eine entsprechende Spezifikation des GARCH-Modells. Da die Anzahl der Parameter von den Spezifikationen beider Modelle gleich ist, liefert das Akaike- und Schwarz-Kriterium das gleiche Ergebnis wie das Kullback-Leibler-Kriterium. Der Vergleich der Student- t -Verteilung mit der N-LN-Mischung bei gegebener Varianzspezifikation zeigt, dass für beide Modelle die Student- t -Verteilung nach dem Informationskriterium von Kullback und Leibler eine bessere Anpassung an die Daten ermöglicht, als die N-LN-Mischung. Somit ist das EGARCH- t -Modell das beste ARCH-Modell gemäß allen drei Informationskriterien für die ausgewählten Aktien.

Letztendlich untersucht Liesenfeld (1998) die Autokorrelation der standardisierten Residuen $\hat{z}_t = (r_t - \hat{\mu}) / \hat{h}_t^{1/2}$ und deren quadrierten Werte \hat{z}_t^2 bei der Laglänge von jeweils $\tau = 20$. Die Nullhypothese der fehlenden Autokorrelation kann weder für \hat{z}_t noch für \hat{z}_t^2 zu den üblichen Signifikanzniveaus verworfen werden. Da die quadrierten Residuen \hat{z}_t^2 in keinem Modell signifikante Autokorrelation aufweisen, sind alle angeführten ARCH-Modelle dazu geeignet, das autoregressive Verhalten der Varianz der Preisänderungen nachzubilden.

Aufgrund der Autokorrelation im nicht beobachtbaren Varianzprozess lässt sich die Likelihoodfunktion für das ARV-Modell nur schwer berechnen. Die Likelihoodfunktion stellt in diesem Fall ein mehrdimensionales Integral dar, dessen Dimension von der Anzahl der Beobachtungen abhängt. Liesenfeld (1998) verwendet deswegen die SML-Methode. Diese Methode basiert auf einer *Importance-Sampling*-Prozedur und hat den Vorteil, dass die Ergebnisse mit den ML-Schätzungen des ARCH-Modells vergleichbar sind.

Liesenfeld (1998) definiert die Likelihoodfunktion des ARV-Modells wie folgt:

$$f(\mathbf{r}; \boldsymbol{\theta}) = \int_{\mathbf{R}^T} f(\mathbf{r}, \boldsymbol{\lambda}; \boldsymbol{\theta}) d\boldsymbol{\lambda} = \int_{\mathbf{R}^T} \prod_{t=1}^T f(r_t | \lambda_t) f(\lambda_t | \lambda_{t-1}) d\boldsymbol{\lambda}. \quad (3.63)$$

Dabei ist $\lambda_t = \ln \sigma_t^2$,

$\boldsymbol{\theta} = (\mu, \phi_0, \phi_1, \phi_2)'$ ist der zu schätzende Parametervektor,

$\mathbf{r} = \{r_t\}_{t=1}^T$ ist der Vektor der beobachtbaren Variablen,

$\boldsymbol{\lambda} = \{\lambda_t\}_{t=1}^T$ ist der Vektor der latenten Variablen.

Die Anzahl der Beobachtungen T bestimmt die Dimension des Integrals. Die von Danielsson und Richard (1993; zit. n. Liesenfeld 1998, S. 63) vorgeschlagene Monte-Carlo (MC) Methode ermöglicht eine nahezu exakte Schätzung dieses hochdimensionalen Integrals $f(\mathbf{r}; \boldsymbol{\theta})$ für einen gegebenen Wert des Parametervektors $\boldsymbol{\theta}$. Um die Genauigkeit der SML-Methode zu überprüfen, ermittelt Liesenfeld (1998) Schätzungen für den MC-Stichprobenfehler der maximierten Log-Likelihoodfunktion. Für alle Datenreihen liegt der Stichprobenfehler bei ungefähr 0,003% des jeweiligen Log-Likelihoodwertes, was eine hinreichende Präzision des SML-Verfahrens bestätigt.

Die SML-Schätzung liefert einen Wert für den Persistenzparameter ϕ_1 , der für alle ausgesuchte Datenreihen signifikant größer als Null ist. Außerdem ist diese Schätzung von der gleichen Größenordnung wie in den Modellen der ARCH-Klasse. Liesenfeld (1998) kommt zum Schluss, dass ARV- und ARCH-Modelle ein ähnlich großes Ausmaß an Persistenz im Varianzprozess prognostizieren.

Das ARV-Modell kann nicht die ganze beobachtbare Leptokurtosis erklären: Der geschätzte Exzess κ_r ist für alle vier Aktien positiv, aber niedriger als die entsprechenden empirischen Werte und ebenfalls niedriger als die ML-Schätzungen aus dem Standardmodell. Liesenfeld (1998) kommt zum Schluss, dass der im dynamischen ARV-Modell erklärte Anteil der Leptokurtosis kleiner ist, als im statischen Mischungsverteilungsmodell.

Gleich wie für die Modelle der ARCH-Klasse, verwendet Liesenfeld (1998) die Log-Likelihoodwerte und die Informationskriterien von Akaike und Schwarz, um die Anpassungsgüte des ARV-Modells mit dem besten Modell der ARCH-Klasse, dem EGARCH- t -Modell, zu vergleichen. Nach allen Kriterien erreicht das EGARCH- t -Modell eine bessere Anpassung an die Daten als das ARV-Modell, wobei zu beachten ist, dass das Letztere einen Parameter weniger besitzt (was für den Wert der Log-Likelihoodfunktion ausschlaggebend ist) und dass der Unterschied zwischen den Modellen relativ gering ist. Das lässt darauf schließen, dass beide Modelle eine ähnlich gute Anpassung ermöglichen.

Der Vergleich der Werte von standardisierten Residuen \hat{z}_t zeigt, dass die Werte aus dem ARV-Modell von der gleichen Größenordnung sind als jene aus den ARCH-Modellen. Wie im Falle der ARCH-Modelle, kann hier die Nullhypothese der Unkorreliertheit der Residuen \hat{z}_t zu den Niveaus von 5% bzw. 10% nicht verworfen werden. Für die quadrierten Residuen \hat{z}_t^2 kann diese Nullhypothese zu keinem der üblichen Signifikanzniveaus abgelehnt werden. Liesenfeld (1998) kommt zum Schluss, dass die geschätzte bedingte Varianz des ARV-Modells die beobachtete Autokorrelation in den quadrierten Preisänderungen gut filtern kann, genauso wie die geschätzte bedingte Varianz der ARCH-Modelle. Auch der Vergleich der Persistenzparameter zeigt, dass die beobachtbare serielle Korrelation in der Varianz der Preisänderungen durch beide Modelle ähnlich gut angepasst werden kann.

3.4 Bivariate Modelle

In seinem Modell führt Clark den signifikant positiven Zusammenhang zwischen der Varianz der Preisänderungen und dem Handelsvolumen darauf zurück, dass das Handelsvolumen positiv mit der Anzahl der preisrelevanten Informationen, welche die Varianz der Preisänderungen bestimmt, korreliert ist. Werden aber die Varianz und das Handelsvolumen gemeinsam von der Informationsrate bestimmt, führt das zu einem Simultanitätsbias im

Standardmodell. Außerdem wird die serielle Korrelation der Informationsrate von Clark nicht berücksichtigt.

In diesem Abschnitt werden bivariate Modelle von Tauchen und Pitts (1983) sowie Andersen (1996) präsentiert. In diesen Modellen wird das Handelsvolumen als eine endogene Variable betrachtet, das heißt, es wird von der Informationsrate gemeinsam mit der Varianz der Preisänderungen bestimmt.

3.4.1 Das Ausgangsmodell

Tauchen und Pitts haben 1983 das Mischungsverteilungsmodell von Clark mit Hilfe eines mikroökonomischen Ansatzes erweitert. Sie verwenden ein Walrasianisches Gleichgewicht, aus dem sie die Verteilung des Handelsvolumens und der Preisänderungen ableiten. In diesem Modell sind die Preisänderungen sowie das Handelsvolumen endogene Variablen, die gemeinsam von der latenten Informationsrate beeinflusst werden, die als Mischungsvariable dient.

Im Modell von Tauchen und Pitts (1983)⁵ wird der Markt mit einem Wertpapier und J aktiven Händlern betrachtet. Während des Tages verläuft der Markt über mehrere verschiedene Walrasianische Gleichgewichte. Der Übergang vom Gleichgewicht $(i-1)$ zum Gleichgewicht i wird durch die Ankunft neuer Information ausgelöst. Die Zeitabschnitte zwischen den Gleichgewichten müssen nicht unbedingt gleich sein. Während des i -ten Gleichgewichts kann die Wunschposition Q_{ij} des j -ten Händlers folgendermaßen beschrieben werden:

$$Q_{ij} = \alpha [P_{ij}^* - P_i] \quad \text{mit} \quad j = 1, 2, \dots, J. \quad (3.64)$$

Dabei ist $\alpha > 0$ eine Konstante, P_{ij}^* ist der Reservationspreis des j -ten Händlers, und P_i ist der aktuelle Preis am Markt. Die Gleichung (3.64) berücksichtigt keine Transaktionskosten, somit unterscheiden sich die Händler nur in der Höhe ihrer Reservationspreise. Ein positiver Wert von Q_{ij} repräsentiert einen Kaufwunsch, ein negativer Wert steht für eine *short* Position. Die Reservationspreise der J aktiven Händler unterscheiden sich vom momentanen Angebotspreis am Markt. Der Unterschied in den Wunschpositionen Q_{ij} resultiert aus den

⁵ Die Notation im Modell wurde tlw. geändert, um konsistent mit Liesenfeld (1998) zu bleiben.

unterschiedlichen Erwartungen über die Zukunft sowie verschiedenen Bedürfnissen zur Risikostreuung. Der Reservationspreis der nicht aktiven Händler ist gleich dem Angebotspreis. Tauchen und Pitts (1983) nehmen an, dass sich die Anzahl der Händler J während eines Tages nicht verändert.

Für ein Gleichgewicht muss $\sum_{j=1}^J Q_{ij} = 0$ gegeben sein. Das impliziert eine Marktträumung beim Durchschnitt der Reservationspreise:

$$P_i = \frac{1}{J} \sum_{j=1}^J P_{ij}^* . \quad (3.65)$$

Als Nächstes betrachten Tauchen und Pitts (1983) den Übergang vom $(i-1)$ -ten zum i -ten Gleichgewicht. Die Ankunft der preisrelevanten Informationen verändert die Reservationspreise der Händler. Die daraus resultierende Änderung im Marktpreis r_i ist der Durchschnitt der Änderungen in den Reservationspreisen. Das Handelsvolumen v_i ist die Hälfte der Summe der absoluten Veränderungen von den Wunschpositionen der Händler. Mit der Gleichgewichtsbedingung und der Gleichung (3.64) können die Preisänderung und das Handelsvolumen folgendermaßen beschrieben werden:

$$r_i = \frac{1}{J} \sum_{j=1}^J \Delta P_{ij}^* , \quad (3.66)$$

$$v_i \equiv \frac{1}{2} \sum_{j=1}^J |Q_{ij} - Q_{i-1,j}| = \frac{\alpha}{2} \sum_{j=1}^J |\Delta P_{ij}^* - r_i| . \quad (3.67)$$

Dabei ist die Veränderung des j -ten Reservationspreises $\Delta P_{ij}^* = P_{ij}^* - P_{i-1,j}^*$. Die Gleichung (3.67) zeigt, dass das Handelsvolumen proportional zur durchschnittlichen absoluten Abweichung der Reservationspreisänderung vom Durchschnitt ist.

Zur Spezifizierung der Wahrscheinlichkeitsverteilung der Änderung der Reservationspreise $\{\Delta P_{ij}^*\}$ zerlegen Tauchen und Pitts (1983) diese Änderung in zwei Komponenten:

$$\Delta P_{ij}^* = \phi_i + \psi_{ij}, \quad (3.68)$$

$$\text{mit } E[\phi_i] = E[\psi_{ij}] = 0, \quad \text{Var}[\phi_i] \equiv \sigma_\phi^2, \quad \text{Var}[\psi_{ij}] \equiv \sigma_\psi^2.$$

Dabei sind ϕ und ψ zeit- und händlerunabhängig. Die Komponente ϕ_i ist für alle Händler gleich, die Komponente ψ_{ij} ist spezifisch für den j -ten Händler. Mit diesen zwei Komponenten kann die Reaktion der Händler auf neue Informationen beschrieben werden. Ist beispielsweise die gemeinsame Komponente ϕ_i groß im Vergleich zur individuellen Komponente ψ_{ij} , reagieren die Händler fast gleich. Im umgekehrten Fall führt die gleiche Information zu eher verschiedenen Reaktionen. Es wird angenommen, dass die Komponenten keine seriellen Abhängigkeiten aufweisen.

Tauchen und Pitts (1983) schreiben die Gleichungen (3.66) und (3.67) unter Verwendung der Komponenten um:

$$r_i = \phi_i + \bar{\psi}_i \quad \text{mit} \quad \bar{\psi}_i \equiv \frac{1}{J} \sum_{j=1}^J \psi_{ij}, \quad (3.69)$$

$$v_i = \frac{\alpha}{2} \sum_{j=1}^J |\psi_{ij} - \bar{\psi}_i|. \quad (3.70)$$

Die Gleichungen (3.69) und (3.70) zeigen, dass die gemeinsame Komponente ϕ_i keine Auswirkungen auf das Handelsvolumen hat. Somit führt ein großer Wert von ϕ_i im Vergleich zu ψ_{ij} zu einer Preisänderung und zu keiner bzw. kleiner Änderung im Handelsvolumen. Die Komponenten der Reservationspreisänderungen sind identisch und unabhängig verteilt; daraus und aus der Gleichung (3.70) folgt, dass das Handelsvolumen *i.i.d.* ist.

Tauchen und Pitts (1983) halten Folgendes fest:

- (i) die Preisänderungen r_i sind normalverteilt;
- (ii) das Handelsvolumen v_i ist für ein großes J annähernd normalverteilt weil $v_i = (\alpha/2) \sum_{j=1}^J |\psi_{ij} - \bar{\psi}_i|$ eine Summe von fast unabhängigen Summanden ist;

(iii) r_i und v_i sind stochastisch unabhängig. Das folgt aus der Unabhängigkeit der Komponenten $r_i = \phi_i + \bar{\psi}_i$ und aus der Abweichung vom Mittelwert, $\psi_{ij} - \bar{\psi}_i$, die in die Gleichung (3.70) des Volumens einfließt.

Tauchen und Pitts (1983) spezifizieren die ersten zwei Momente für r_i und v_i :

$$\mu_r \equiv E[r_i] = 0, \quad (3.71)$$

$$\sigma_r^2 \equiv \text{Var}[r_i] = \sigma_\phi^2 + \frac{\sigma_\psi^2}{J}, \quad (3.72)$$

$$\mu_v \equiv E[v_i] = \left(\frac{\alpha}{2}\right) \sigma_\psi \sqrt{\frac{2}{\pi}} \left(\sqrt{\frac{J-1}{J}}\right) J, \quad (3.73)$$

$$\sigma_v^2 \equiv \text{Var}[v_i] = \left(\frac{\alpha}{2}\right)^2 \sigma_\psi^2 \left(1 - \frac{2}{\pi}\right) J. \quad (3.74)$$

Die Zufallsvariablen r_i und v_i sind zwar stochastisch unabhängig, aber deren gemeinsame Abhängigkeit von den Komponenten ϕ_i und ψ_{ij} begründet Abhängigkeiten zwischen deren Momenten.

Die Anzahl J der Händler ist nicht zufällig, sie ist für jeden Tag fixiert. Die Anzahl der täglichen Gleichgewichte wird durch die zufällige Anzahl der eingetroffenen preisrelevanten Informationen n_t bestimmt. Die täglichen Preisänderungen r_t ergeben sich aus der Summe der Teilpreisänderungen r_i , das tägliche Handelsvolumen v_t ergibt sich aus der Summe der Transaktionsvolumina v_i . Die Anzahl der Summanden variiert von Tag zu Tag und hängt von n_t ab.

Liesenfeld (1998) fasst dieses Ausgangsmodell wie folgt zusammen:

$$r_t | n_t \sim N(0, \sigma_r^2 n_t) \quad (3.75)$$

$$v_t | n_t \sim N(\mu_v n_t, \sigma_v^2 n_t) \quad (3.76)$$

mit $\text{Cov}(r_t, v_t | n_t) = 0$

Die gemeinsame Verteilung der Preisänderungen und des Handelsvolumens ist eine Mischung von bivariaten Normalverteilungen, wobei die latente Informationsrate die gemeinsame Mischungsvariable darstellt. Der Unterschied zum Standardmodell von Clark liegt also darin, dass das Handelsvolumen als eine endogene Variable behandelt wird.

In diesem Modell sind die Preisänderungen und das Handelsvolumen unkorreliert.

$$\text{Cov}(r_t, v_t) = 0 \quad (3.77)$$

Tauchen und Pitts (1983) schreiben das Modell unter Verwendung von den Gleichungen (3.75) und (3.76) um:

$$r = \sigma_r \sqrt{n_t} Z_1, \quad (3.78)$$

$$v = \mu_v n_t + \sigma_v \sqrt{n_t} Z_2, \quad (3.79)$$

wobei Z_1 und Z_2 Zufallsvariablen mit $N(0,1)$ sind. Außerdem sind Z_1 , Z_2 und n_t von einander unabhängig. Jetzt kann das Verhältnis zwischen den quadrierten Preisänderungen und dem Handelsvolumen untersucht werden:

$$\begin{aligned} \text{Cov}(r^2, v) &= E(r^2 v) - E(r^2)E(v) = \sigma_r^2 \mu_v E(n_t^2) - \sigma_r^2 \mu_v (E[n_t])^2 \\ &= \sigma_r^2 \mu_v \text{Var}(n_t) > 0 \end{aligned} \quad (3.80)$$

Für die quadrierten Preisänderungen und das Handelsvolumen ergibt sich also eine positive Korrelation. Dieser Zusammenhang ist darauf zurückzuführen, dass beide Variablen – v_t und die Varianz von r_t^2 – positiv von der dritten Variable – der Informationsrate n_t – abhängen. Somit liefert das Modell eine Erklärung für den positiven Zusammenhang zwischen der Varianz der Preisänderungen und dem Handelsvolumen.

Da im statischen bivariaten Modell sowohl die bedingte Varianz der Preisänderungen als auch die ersten beiden bedingten Momente des Handelsvolumens von der latenten Informationsrate abhängen, sieht Liesenfeld (1998) eine Möglichkeit, aus dem Handelsvolumen zusätzliche Informationen über die latente Größe zu gewinnen.

3.4.2 Das modifizierte Modell

Andersen hat 1996 das statische bivariate Modell (3.75) und (3.76) von Tauchen und Pitts (1983) modifiziert. Er verwendet ein Mikrostrukturmodell von Glosten und Milgrom (1985; zit. n. Andersen 1996, S. 171 ff.) zur Herleitung seines dynamischen bivariaten Mischungsverteilungsmodells.

Andersen (1996) beschreibt das Mikrostrukturmodell wie folgt. Es gibt ein Wertpapier und drei Gruppen von risikoneutralen Marktteilnehmern. Der Marketmaker, ein Spezialist, bildet die erste Gruppe. Der Marketmaker ist bereit, das Wertpapier zu dem von ihm bestimmten Preis jederzeit zu kaufen oder zu verkaufen. Der Preis wird von ihm so festgelegt, dass sein erwarteter Gewinn gleich Null ist. Insider bilden die zweite Gruppe. Das sind informierte Händler, die stochastische private Informationen über den zukünftigen Wert des Wertpapiers erhalten. Insider versuchen, diesen Informationsvorsprung gegenüber den anderen Marktteilnehmern auszunutzen. Die dritte Gruppe sind nichtinformierte Liquiditätshändler, die aus modellexogenen Gründen das Wertpapier kaufen bzw. verkaufen. Investoren treffen sequentiell, zufällig und anonym auf den Markt ein und entscheiden, ob sie das Wertpapier zum Preis vom Spezialisten kaufen bzw. verkaufen wollen oder überhaupt nicht handeln. Insider erhalten preisrelevante Informationen während des Handelstages. Wichtig ist, dass eine Reihe von abgeschlossenen Geschäften nach relativ kurzer Zeit diese privaten Informationen aufdeckt. Im Endeffekt werden sich alle Marktteilnehmer über den Wert des Wertpapiers einig, bis zum Eintreffen der nächsten privaten Information. Das heißt, dass die Ankunft der privaten Information einen dynamischen Lernprozess auslöst, der in Preisen resultiert, die den Inhalt dieser Information voll widerspiegeln. Das Eintreffen der nicht informierten Händler auf den Markt folgt einem Poisson Prozess mit einer konstanten Intensität m_0 am Tag. Das minimiert den Einfluss der Liquiditätshändler auf die Dynamik der Renditenvolatilität und des Handelsvolumens. Es wird angenommen, dass Liquiditätshändler eine unelastische Nachfrage- bzw. Angebotsfunktion haben, somit beträgt die

Wahrscheinlichkeit für deren Kauf bzw. Verkauf 0,5. Im Gegenteil dazu basiert die Handelsentscheidung eines Insiders auf dem Wert des Wertpapiers, der von seiner Informationsausstattung abhängt. Insider bekommen ähnliche, aber nicht gleiche Informationen, das heißt, deren Einschätzung des Wertes kann unterschiedlich sein. Somit hängt das Eintreffen der Insider von deren privaten Signalen sowie von deren Handelsgeschichte, *bid-ask spread* und anderen Faktoren ab.

Andersen (1996) unterstreicht folgende Implikationen dieses Mikrostrukturmodells. Einerseits folgen die Transaktionspreise einem Martingal, das heißt, dass der Erwartungswert gegeben dem Informationsstand des Marketmakers Null ist. Andererseits löst die Ankunft der privaten Informationen einen Verarbeitungsprozess aus, der in einem Marktgleichgewicht endet. Dieses Marktgleichgewicht ist dadurch charakterisiert, dass sich alle Marktteilnehmer über den Preis des Wertpapiers einig sind und dass dieser Preis die privaten Informationen vollständig widerspiegelt.

Andersen (1996) unterstellt, dass die Anzahl n_t ⁶ der privaten Informationen am Tag eine stochastische Größe ist. Es wird angenommen, dass die Änderung der Gleichgewichtspreise und die dazu gehörigen Transaktionsvolumina *i.i.d.* Zufallsvariablen sind. Die tägliche Preisänderung r_t ergibt sich als eine Summe der Gleichgewichtspreisänderungen mit der Anzahl der Summanden n_t . Für hinreichend große n_t lässt sich für dieses Modell eine bedingte Verteilung von r_t ableiten, die der Spezifikation von Clark (1973) und Tauchen und Pitts (1983) entspricht, vgl. (3.5) und (3.75):

$$r_t | n_t \sim N(0, \sigma_r^2 n_t).$$

Andersen (1996) unterteilt das Handelsvolumen in eine informationsbedingte und eine *noise*-Komponente, die durch Liquiditätshändler verursacht wird: $v_t = v_t^i + v_t^l$. Die Volumenskomponente der Liquiditätshändler v_t^l wird durch einen stochastischen Prozess mit einer konstanten Intensität m_0 bestimmt. Die Wahrscheinlichkeit $g_{t,i}$, mit welcher ein Insider während der i -ten Informationsverarbeitungsphase des Tages t eine Transaktion durchführt, ist laut Liesengeld (1998) für alle i und t klein und folgt einem stationären Prozess. Andersen

⁶ Die Notation im Modell wurde tlw. geändert, um konsistent mit Liesenfeld (1998) zu bleiben.

(1996) nimmt an, dass die Anzahl der Insider F groß ist. Aus diesen Annahmen folgt für das Handelsvolumen der Liquiditätshändler eine Poissonverteilung:

$$v_t^l \sim \text{Po}(m_0) \quad \text{mit} \quad m_0 > 0 \quad (3.81)$$

Für die Bestimmung der Verteilungsannahme für das tägliche Handelsvolumen der Insider betrachtet Andersen (1996) zunächst das Transaktionsvolumen $\sum_{j=1}^F t_{t,i,j}$, das durch die Ankunft der i -ten Information ausgelöst wird. Mit $t_{t,i,j}$ wird eine Indikatorvariable bezeichnet. Sie gibt an, ob der j -ter Insider aufgrund der i -ten Information des Tages t eine Transaktion durchführt ($t_{t,i,j} = 1$) oder nicht ($t_{t,i,j} = 0$). Die Indikationsvariable $t_{t,i,j}$ folgt einer Bernoulli-Verteilung mit Parameter $g_{t,i}$. Daraus folgt, dass die Summe der Indikationsvariablen $\sum_{j=1}^F t_{t,i,j}$ binomial verteilt mit Parametern $g_{t,i}$ und F ist. Wenn die Anzahl der Insider F groß und die Wahrscheinlichkeit $g_{t,i}$ klein ist, lässt sich eine Binomialverteilung durch eine Poissonverteilung mit Parameter $F \cdot g_{t,i}$ sehr gut approximieren. Für das Transaktionsvolumen ergibt sich also:

$$\sum_{j=1}^F t_{t,i,j} \sim \text{Po}(F \cdot g_{t,i}). \quad (3.82)$$

Das tägliche Handelsvolumen aus den Transaktionen der Insider v_t^i ergibt sich also als eine Summe der unabhängig poissonverteilten Zufallsvariablen. Liesenfeld (1998) spezifiziert die bedingte Verteilung von v_t^i gegeben die Informationsrate n_t wie folgt:

$$v_t^i | n_t \sim \text{Po}\left(\sum_{i=1}^{n_t} F \cdot g_{t,i}\right). \quad (3.83)$$

Er nimmt an, dass die Wahrscheinlichkeit $g_{t,i}$ durch einen stationären Prozess erzeugt wird und dass n_t groß ist. In diesem Fall kann der Parameter der Poissonverteilung in (3.83) folgendermaßen approximiert werden:

$$F \sum_{i=1}^{n_t} g_{t,i} = F n_t \left(\frac{1}{n_t} \sum_{i=1}^{n_t} g_{t,i} \right) \approx F n_t E(g_{t,i}). \quad (3.84)$$

Setzt man $m_1 = F \cdot E(g_{t,i})$, ergibt sich aus (3.83) und (3.84):

$$v_t^i | n_t \sim \text{Po}(m_1 n_t) \quad \text{mit} \quad m_1 \geq 0. \quad (3.85)$$

Die Variable m_1 kann als ein Faktor der Sensibilität interpretiert werden, das heißt, wie stark das Handelsvolumen auf die preisrelevanten Informationen reagiert.

Liesenfeld (1998) fasst das modifizierte Modell von Andersen (1996) wie folgt zusammen:

$$r_t | n_t \sim N(0, \sigma_r^2 n_t), \quad (3.86)$$

$$v_t | n_t \sim \text{Po}(m_0 + m_1 n_t), \quad (3.87)$$

$$\text{mit} \quad \text{Cov}(r_t, v_t | n_t) = 0.$$

Das modifizierte Modell unterscheidet sich vom statischen Modell (3.75) und (3.76) dadurch, dass das Handelsvolumen poissonverteilt ist. Die Normalverteilung der Preisänderungen ist gleich in beiden Modellen. Außerdem wird im Modell von Andersen (1996) das Handelsvolumen in zwei Komponenten unterteilt. Dabei ist $m_1 n_t$ das informationsinduzierte Handelsvolumen. Die Komponente m_0 ist von der Ankunft der preisrelevanten Informationen unabhängig.

Bisher wurde die Stationarität der Handelsvolumenzeitreihen angenommen. Das steht im Gegensatz dazu, dass das Handelsvolumen typischerweise einen positiven langfristigen Trend aufweist und somit nicht stationär ist. Andersen (1996) schlägt eine Bereinigung der Daten um den Trend π_t vor. Das beobachtbare Handelsvolumen ist somit $v_t^0 = v_t \cdot \pi_t$, dabei repräsentiert v_t das stationäre Handelsvolumen wie oben angeführt (3.87). Der Wert des unbedingten Mittelwertes der Reihe wird in diesem Modell aber nicht definiert. Das heißt, dass der geschätzte Trend $\hat{\pi}_t$ im besten Fall einem c -fachen des echten Trends entspricht: $c \cdot \hat{\pi}_t = \pi_t$. Dabei ist c eine positive unbekannte Konstante, die gemeinsam mit den anderen

Parametern des Modells geschätzt werden muss. Das trendbereinigte Volumen ergibt sich als $\hat{v}_t = v_t^0 / \hat{\pi}_t = cv_t$. Somit ist die bedingte Verteilung für das trendbereinigte Handelsvolumen folgendermaßen gegeben:

$$\hat{v}_t | n_t \sim c \cdot \text{Po}(m_0 + m_1 n_t) \quad \text{mit} \quad c > 0. \quad (3.88)$$

3.4.3 Schätzung der Modelle und Vergleich der empirischen Ergebnisse

Zur Schätzung des modifizierten Modells verwendet Liesenfeld (1998) auf der Likelihoodfunktion basierende Verfahren, welche die vollständige Spezifikation der bedingten Dichtefunktion berücksichtigen. Die bedingte Dichte für \hat{v}_t / c ist hier nur für ganzzahlige Zufallsvariablen definiert. Um das zu gewährleisten, nimmt Liesenfeld (1998) an, dass $(m_0 + m_1 n_t)$ bedingt durch eine große Anzahl der Informationen hinreichend groß ist, und approximiert die Poissonverteilung durch folgende Normalverteilung:

$$\hat{v}_t | n_t \sim N(c[m_0 + m_1 n_t], c^2[m_0 + m_1 n_t]). \quad (3.89)$$

Liesenfeld (1998) verwendet im Folgenden diese Approximation für Schätzung des Modells von Andersen (1996).

Um die Spezifikation der bivariaten Modelle zu vervollständigen, trifft Liesenfeld (1998) Annahmen über das stochastische Verhalten der Informationsrate.

Im Abschnitt 3.3 wurde erläutert, dass sowohl das ARV-Modell als auch die ARCH-Modelle eine serielle Korrelation in der Informationsrate implizieren. Somit eignen sich beide Ansätze zur Modellierung eines dynamischen bivariaten Ansatzes. Liesenfeld (1998) verwendet dafür die sich aus dem ARV-Modell ergebende Zeitreihenspezifikation der Informationsrate, weil sie keine verzögerten endogenen Variablen enthält (im Gegensatz zum ARCH-Ansatz):

$$\ln n_t = \phi_0 + \phi_1 \ln n_{t-1} + \phi_2 \eta_t \quad \text{mit} \quad \eta_t \sim i.i.d.N(0,1) \quad (3.90)$$

Das ist ein AR(1)-Prozess für die logarithmierte Informationsrate mit normalverteilten Innovationen und Persistenzparameter ϕ_1 . Das Modell von Tauchen und Pitts (1983) ist ein statischer Spezialfall dieser Spezifikation mit $\phi_1 = 0$.

Aus den Gleichungen (3.90), (3.75) und (3.76) leitet Liesenfeld (1998) folgende Spezifikation für die täglichen Preisänderungen und das trendbereinigte Handelsvolumen im statischen Modell (dabei ist $\lambda_t = \ln n_t$) her:

$$r_t | \lambda_t \sim N(\mu_r, \sigma_r^2 e^{\lambda_t}), \quad (3.91)$$

$$\hat{v}_t | \lambda_t \sim N(\mu_v e^{\lambda_t}, \sigma_v^2 e^{\lambda_t}), \quad (3.92)$$

$$\text{mit } \lambda_t | \lambda_{t-1} \sim N([\phi_0 + \phi_1 \lambda_{t-1}], \phi_2^2). \quad (3.93)$$

Für das dynamische Modell wird die Gleichung (3.92) gemäß (3.87) und (3.89) durch

$$\hat{v}_t | \lambda_t \sim N(c[m_0 + m_1 e^{\lambda_t}], c^2[m_0 + m_1 e^{\lambda_t}]) \quad (3.94)$$

ersetzt. Sollten die Spezifikationen (3.91)-(3.94) dazu geeignet sein, die gemeinsame Dynamik der Preisänderungen und des Handelsvolumens zu beschreiben, wird die latente Informationsrate $n_t = e^{\lambda_t}$ die beobachtbare serielle Korrelation im Varianzprozess der Preisänderungen erfassen. In diesem Fall kann aus dem Handelsvolumen zusätzliche Information über den Prozess, der die Dynamik der Varianz der Preisänderungen bestimmt, gewonnen werden.

Aufgrund der Autokorrelation in den Spezifikationen der bivariaten Modelle (3.91)-(3.93) und (3.91), (3.93) und (3.94) ergeben sich die Likelihoodfunktionen analog zum ARV-Modell als hochdimensionale Integrale. Deswegen hat Andersen (1996) auf die ML-Methode verzichtet und sein dynamisches bivariates Modell mit Hilfe der GMM geschätzt. Liesenfeld (1998) verwendet hier noch mal die SML-Methode, da sie effizienter als die GMM ist und die Verwendung der herkömmlichen Instrumente für die ML-Methode ermöglicht.

Um die Likelihoodfunktion für die bivariaten Modelle zu erhalten, integriert Liesenfeld (1998) den Vektor der latenten Variablen $\lambda = \{\lambda_t\}_{t=1}^T$ aus der jeweiligen gemeinsamen Dichtefunktion von $\mathbf{r} = \{r_t\}_{t=1}^T$, $\mathbf{v} = \{v_t\}_{t=1}^T$ und λ heraus. Damit sind die Likelihoodfunktionen durch folgendes T -dimensionale Integral gegeben:

$$f(\mathbf{r}, \mathbf{v}; \boldsymbol{\theta}) = \int_{\mathbf{R}^T} f(\mathbf{r}, \mathbf{v}, \boldsymbol{\lambda}; \boldsymbol{\theta}) d\boldsymbol{\lambda}. \quad (3.95)$$

Der zu schätzende Parametervektor $\boldsymbol{\theta}$ ist für das statische Modell mit $(\mu_r, \mu_v, \sigma_v, \phi_0, \phi_1, \phi_2)'$ und für das modifizierte Modell mit $(\mu_r, c, m_0, m_1, \phi_0, \phi_1, \phi_2)'$ gegeben. Liesenfeld (1998) verwendet die ML-Methode für die Schätzung der statischen Spezifikationen und die SML-Methode – wie bei dem ARV-Modell im Abschnitt 3.3.3 beschrieben – für die dynamischen Spezifikationen.

Zuerst schätzt Liesenfeld (1998) Parameter der statischen Spezifikation der Modelle von Tauchen und Pitts (1983) und Andersen (1996), das heißt, der Persistenzparameter ϕ_1 in der Gleichung (3.93) ist gleich Null.

Die ML-Schätzungen des statischen Ausgangsmodells (3.91)-(3.93) zeigen, dass der Variationsparameter der Mischungsvariable ϕ_2 signifikant größer als Null ist und zwischen 0,37 und 0,45 liegt. Die Schätzungen für den Exzess κ_τ sind ebenfalls signifikant größer als Null aber deutlich kleiner als die empirischen Werte. Das heißt, dass nur ein kleiner Teil der beobachteten Leptokurtosis durch dieses Modell erklärt werden kann.

Das modifizierte Modell von Andersen (1996) (3.91), (3.93), (3.94) kann bei $\phi_1 = 0$ als eine Verallgemeinerung des Ausgangsmodells betrachtet werden, da es einen von der Ankunft der Information unabhängigen Teil des Handelsvolumens berücksichtigt. Dieser Teil wird durch den Parameter $c \cdot m_0$ erfasst, der für alle getesteten Aktien signifikant größer als Null geschätzt wird. Dabei beträgt dieser von der Informationsankunft unabhängige Volumen zwischen 39% und 66%. Die Schätzungen für den Parameter ϕ_2 sind hier ebenfalls signifikant größer als Null und mit Werten von 0,67 bis 0,95 deutlich größer als im Modell von Tauchen

und Pitts (1983). Die Schätzungen für den Exzess κ_τ sind ebenfalls deutlich höher; für eine der Aktien wird die empirische Leptokurtosis durch das modifizierte Modell sogar vollständig erklärt.

Zum Vergleich der statischen Spezifikationen wurde das Informationskriterium von Schwarz herangezogen. Der Wert für alle vier Aktien ist im modifizierten Modell deutlich geringer, somit ermöglicht dieses Modell eine bessere Anpassung an die Daten.

Des Weiteren schätzt Liesenfeld (1998) mithilfe der SML-Methode die Parameter der dynamischen Spezifikationen, welche die Autokorrelation der Informationsrate mit $\phi_1 \neq 0$ zulassen und eine bivariate Erweiterung des ARV-Modells darstellen.

Wie schon für das ARV-Modell, ermittelt Liesenfeld (1998) für das dynamische Modell (3.91)-(3.93) Schätzungen für den MC-Stichprobenfehler der maximierten Log-Likelihoodfunktion. Er kommt zum Ergebnis, dass die MC-Stichprobenfehler im bivariaten Modell sehr klein sind, sogar signifikant kleiner als die im ARV-Modell. Das bedeutet, dass sich das dynamische bivariate Modell mit der SML-Methode sehr genau schätzen lässt. Die SML-Schätzungen für den Persistenzparameter ϕ_1 sind alle signifikant größer als Null und liegen zwischen 0,60 und 0,65. Um die dynamische Spezifikation des Ausgangsmodells mit der statischen Spezifikation zu vergleichen, berechnet Liesenfeld (1998) die Likelihood-Ratio-Statistik $LR_{\phi_1=0}$ der Nullhypothese $H_0 : \phi_1 = 0$. Er kommt zum Ergebnis, dass die dynamische Spezifikation gegenüber der statischen eine erhebliche Verbesserung darstellt. Bei einem Vergleich der Schätzungen des Persistenzparameters ϕ_1 im bivariaten Modell und im univariaten ARV-Modell wird ersichtlich, dass die Schätzungen im Ausgangsmodell deutlich kleiner sind als jene im ARV-Modell, siehe dazu Tabelle 3. Schließlich ist an den geringen Werten für den Exzess κ_τ zu erkennen, dass das Modell nur einen kleinen Teil der empirischen Leptokurtosis erklärt.

Des Weiteren wird die dynamische Spezifikation des modifizierten Modells (3.91), (3.93), (3.94) geschätzt. Die geschätzten MC-Stichprobenfehler der Log-Likelihoodfunktionen sind zwar nicht so klein wie beim Modell von Tauchen und Pitts (1983), aber gering genug, damit die Schätzung des Modells als genau genug betrachtet werden kann. Der Parameter $c \cdot m_0$ ist in der dynamischen Spezifikation ebenso signifikant größer als Null, der von der Ankunft der

Information unabhängige Teil des Handelsvolumens wird zwischen 13,7% und 30,5% prognostiziert. Bei einem Vergleich der dynamischen Spezifikation des modifizierten Modells mit der statischen Spezifikation kommt Liesenfeld (1998) zum Ergebnis, dass die Berücksichtigung der Autokorrelation in der Informationsrate eine signifikante Verbesserung darstellt: Die Werte der Likelihood-Ratio-Statistik $LR_{\phi_1=0}$ zeigen an, dass die Nullhypothese $H_0 : \phi_1 = 0$ zu jedem üblichen Signifikanzniveau verworfen wird. Die Schätzungen für den Persistenzparameter ϕ_1 sind zwar größer, als in der dynamischen Spezifikation des Modells von Tauchen und Pitts (1983), aber noch immer deutlich kleiner als die Werte aus dem ARV-Modell (Tabelle 3). Die Schätzungen für den Exzess κ_τ sind größer als in der dynamischen Spezifikation des Ausgangsmodells, das heißt, dass das dynamische modifizierte Modell einen größeren Teil der empirischen Kurtosis erklären kann. Die Werte des Schwarz-Kriteriums zeigen, dass das modifizierte Modell für drei aus vier untersuchten Aktien die bevorzugte Spezifikation darstellt.

<i>Modell</i>	<i>Siemens</i>	<i>Daimler-Benz</i>	<i>Volkswagen</i>	<i>Deutsche Bank</i>
ARV	0,953 (0,017)	0,954 (0,014)	0,906 (0,029)	0,963 (0,013)
Tauchen und Pitts (1983)	0,599 (0,029)	0,654 (0,028)	0,612 (0,021)	0,601 (0,013)
Andersen (1996)	0,658 (0,030)	0,704 (0,031)	0,654 (0,028)	0,706 (0,028)

In Klammern sind die Standardfehler angegeben.

Tabelle 3: Schätzungen des Persistenzparameters.

Quelle: Liesenfeld (1998), Seite 111.

Zusammenfassend stellt Liesenfeld (1998) fest, dass die dynamische bivariaten Spezifikationen eine signifikante Verbesserung gegenüber den statischen Spezifikationen darstellen. Dabei passt sich das Modell von Andersen (1996) in der Regel besser an die Daten an und erklärt einen größeren Teil der empirischen Leptokurtosis als das Modell von Tauchen und Pitts (1983). Die Untersuchung der dynamischen bivariaten Modelle hat gezeigt, dass die Berücksichtigung des Handelsvolumens als eine endogene Variable zu einer deutlichen Reduzierung des Persistenzparameters ϕ_1 im Vergleich zum univariaten ARV-Modell führt

(Tabelle 3). Liesenfeld (1998) führt diese Tatsache darauf zurück, dass es noch eine Quelle der Persistenz in der Varianz der Preisänderungen gibt.

Um dies zu überprüfen, erweitert Liesenfeld (1998) die bivariaten Mischungsverteilungsmodelle (3.91)-(3.94) um eine separate Persistenzquelle zusätzlich zur Persistenz, die durch die gemeinsame Mischungsvariable erfasst wird. Dazu verallgemeinert er die Spezifikation der Preisänderungen (3.91) folgendermaßen:

$$r_t | \lambda_t \sim N(\mu_r, \sigma_r^2 e^{\lambda_t + \gamma_t}) \quad (3.96)$$

$$\text{mit } \gamma_t = \sum_{j=1}^k \alpha_j |r_{t-j} - \mu_r|.$$

Dabei bleiben die Gleichungen für das Handelsvolumen \hat{v}_t und die Mischungsvariable λ_t gleich. Liesenfeld (1998) spezifiziert den Term γ_t als eine Funktion der verzögerten absoluten Residuen der Preisänderungen $|r_{t-j} - \mu_r|$, welche die verzögerten Varianzen der Preisänderungen messen. Sollte die ganze im Varianzprozess vorhandene Persistenz durch die gemeinsame Mischungsvariable erklärt werden, so sollten sich die Koeffizienten $\alpha_j, j = 1, \dots, k$ nicht signifikant von Null unterscheiden, was bedeuten würde, dass die absoluten verzögerten Residuen keinen Erklärungsgehalt besitzen. Um das zu überprüfen, schätzt Liesenfeld (1998) die bivariaten Modelle mit der verallgemeinerten Spezifikation der Preisänderungen (3.96) mit der SML-Methode. Basierend auf dieser Schätzung wird die Nullhypothese $H_0 : \alpha_j = 0$ mit der Likelihood-Ratio-Statistik $LR_{\alpha_j=0}$ getestet.

Die SML-Schätzungen des verallgemeinerten Ausgangsmodells (3.92), (3.93), (3.96) zeigen, dass die Koeffizienten α_j in allen Fällen signifikant größer als Null sind. Außerdem zeigen die Werte der Likelihood-Ratio-Statistik $LR_{\alpha_j=0}$, dass die verzögerten absoluten Residuen einen Erklärungsgehalt besitzen und dass die Nullhypothese zu jedem üblichen Signifikanzniveau abgelehnt wird. Die SML-Schätzung des verallgemeinerten modifizierten Modells (3.93), (3.94), (3.96) führen zu ähnlichen Ergebnissen.

Dieser Test bringt Liesenfeld (1998) zum Schluss, dass die dynamischen bivariaten Mischungsverteilungsmodelle die Persistenz im Varianzprozess der Preisänderungen nicht

vollständig erfassen können. Das steht im Gegensatz zu Restriktion der bivariaten Mischungsverteilungsmodelle, wonach das dynamische Verhalten der Varianz der Preisänderungen ausschließlich durch das Zeitreihenverhalten der gemeinsamen Mischungsvariable n_t bestimmt wird.

Andersen (1996) benutzt die GMM zur Schätzung seines modifizierten Modells und kommt zu ähnlichen Ergebnissen: Der Persistenzparameter im bivariaten Modell ist signifikant kleiner als der im univariaten Modell. Er interpretiert dieses empirische Ergebnis derart, dass der Varianzprozess der Preisänderungen und das Handelsvolumen nicht nur durch einen Informationsprozess, sondern durch zwei oder mehrere Informationsprozesse bestimmt werden, die den Varianzprozess und das Handelsvolumen auf verschiedene Arten beeinflussen. Andersen (1996) nennt periodische makroökonomische Bekanntgaben und sog. „Hexentage“, an denen mehrere Options- und Terminkontrakte gleichzeitig fällig werden, als Beispiel für Ereignisse, die ein stark erhöhtes Handelsvolumen bei einer kurzfristig leicht erhöhten Persistenz im Varianzprozess hervorrufen. Liesenfeld (1998) betont, dass es auch solche Informationen gibt, die eine lange anhaltende Wirkung auf die Varianz der Preisänderungen besitzen, aber das Handelsvolumen fast nicht beeinflussen. Er bringt als Beispiel Gerüchte über Übernahmeabsichten, die sich länger anhalten. Nach seinen empirischen Untersuchung vermutet Liesenfeld (1998), dass das Handelsvolumen vor allem Aufschluss über jenen Informationsprozess gibt, der nur kurzfristige, transitorische Auswirkungen auf den Varianzprozess hat und eine geringe Persistenz erzeugt.

3.5 Zusammenfassung

In diesem Kapitel wurden das Standardmodell von Clark und seine Erweiterungen behandelt. Zuerst wurde die von Clark erarbeitete Mischungsverteilungshypothese präsentiert. Das ist ein univariates Modell, in dem die Volatilität der Preisänderungen durch die Informationsrate bestimmt wird. Da die Anzahl der Informationen eine latente Variable ist, wird dafür das Handelsvolumen als eine Proxyvariable verwendet. In diesem Modell wird das Handelsvolumen als exogen betrachtet und die beobachtete Autokorrelation der Informationsrate nicht berücksichtigt.

Der ARCH-Ansatz von Bollerslev et al. (1994) und der ARV-Ansatz von Taylor (1986) stellen eine dynamische Erweiterung des Mischungsverteilungsmodells von Clark dar. Diese Modelle berücksichtigen die Autokorrelation in der Varianz und sind auch mit der leptokurtischen Gestalt der Preisänderungen konsistent.

Einen anderen Ansatz liefern Tauchen und Pitts (1983). In ihrem bivariaten Modell wird die Varianz der Preisänderungen gemeinsam mit dem Handelsvolumen von der latenten Informationsrate beeinflusst. Sie verwenden ein Walrasianisches Gleichgewicht, um die Verteilung der Preisänderungen und des Handelsvolumens abzuleiten. Allerdings lassen sie keine serielle Abhängigkeit in der Informationsrate zu.

Anders im bivariaten Mischungsverteilungsmodell von Andersen (1996). Dieses Modell berücksichtigt auch die dynamischen Eigenschaften der Prozesse, indem die serielle Korrelation der Informationsrate zugelassen wird. Im Modell wird das Handelsvolumen als eine endogene Variable betrachtet, das heißt das Volumen wird gemeinsam mit der Varianz der Preisänderungen vom Informationsprozess beeinflusst. Außerdem wird der Teil des Handelsvolumens berücksichtigt, der von der Ankunft der Information unabhängig ist. Die empirische Überprüfung dieses Modells zeigt, dass die Berücksichtigung der Autokorrelation in der Informationsrate eine signifikante Verbesserung darstellt und dass sich dieses Modell besser an Daten anpasst. Allerdings zeigt sich auch, dass die Berücksichtigung des Handelsvolumens als eine endogene Variable zu einer deutlichen Reduzierung des Persistenzparameters im Vergleich zum univariaten ARV-Modell führt.

Im nächsten Kapitel werden die möglichen Ursachen für dieses Verhalten des Persistenzparameters untersucht.

4. Ansätze zur Erklärung des Verhaltens des Persistenzparameters

4.1 Einführung

In der Fachliteratur gibt es verschiedene Ansätze, die den von Liesenfeld und Andersen beobachteten Rückgang der geschätzten Persistenzparameter in bivariaten Modellen erklären könnten. In diesem Kapitel werden die möglichen Gründe hierfür präsentiert.

In den Arbeiten von Kumar (2008) und Fleming et al. (2006) wird die Vermutung von Liesenfeld (1998) und Andersen (1996), nach welcher der Informationsfluss aus mehreren heterogenen Komponenten besteht, untersucht. Kumar (2008) erklärt die unterschiedlichen Einflussgrößen auf das Handelsvolumen und somit indirekt die Heterogenität der Informationsrate mit den Ansätzen aus der Portfoliotheorie und des *behavioural finance*. Fleming et al. (2006) verallgemeinern das Modell von Andersen (1996), indem sie auf die reine AR(1)-Spezifikation der Informationsrate verzichten und eine persistente sowie eine nicht persistente Komponente im Informationsfluss zulassen.

In seiner Arbeit erweitert Liesenfeld (2001) das Modell von Tauchen und Pitts (1983) um einen zusätzlichen Faktor, nämlich die Sensibilität der Marktteilnehmer, mit der sie auf die Informationen reagieren. Liesenfeld (2001) kommt zum Schluss, dass die *long-run* Komponente der Volatilität mit dem Prozess der Informationssensitivität der Händler verbunden ist, während die Handelsvolumenzeitreihen die *short-run* Bewegungen der Volatilität „herausfiltern“.

Mestel (2008) untersucht die Preisänderungs- und die Handelsvolumenzeitreihen und findet heraus, dass das *short-* und *long-memory* Verhalten dieser Zeitreihen unterschiedlich ist. Im Handelsvolumen ist die Information über die transitorische Komponente des Varianzprozesses enthalten, während die Ereignisse mit persistenter Wirkung auf das Handelsvolumen und die Renditevarianz unterschiedlichen Einfluss haben. Ein Vergleich der Differenzparameter zeigt, dass die Handelsvolumenzeitreihen eine andere langfristige Dynamik aufweisen und somit für die Erklärung der Preisänderungsreihe nur bedingt geeignet sind.

4.2 Heterogene Komponenten in der Informationsrate

Kumar (2008) bietet eine qualitative Unterscheidung der Faktoren, die auf das Handelsvolumen Einfluss nehmen und heterogene Komponenten des Informationsflusses bilden können. Diese Faktoren liegen einerseits in der Portfoliotheorie und andererseits in der begrenzten Rationalität der Investoren.

Die Volatilität des Handelsvolumens ist über die Zeit nicht konstant. Sie wird von firmenspezifischen Faktoren wie Renditen, systematisches und unsystematisches Risiko, historische Preise, Eigentumsverhältnisse, Marktkapitalisierung und Transaktionskosten beeinflusst. Zu den behaviouristischen Faktoren zählen unterschiedliche Überzeugungen hinsichtlich der Preisentwicklung, psychologische Verzerrungen wie Dispositionseffekt, *tax-loss-selling* Effekt und der Umstand, ob Optionen auf eine Aktie gehandelt werden oder nicht.

Somit fließen verschiedenste Informationen in das Handelsvolumen ein. Die Dynamik des Handelsvolumens kann teilweise durch die „alten“, allgemein bekannten Informationen wie Branchenzugehörigkeit, Änderungen am Aktienmarkt und Verhaltensmuster erklärt werden. Nur ein Teil der Dynamik kann auf das Eintreffen neuer preisrelevanter Informationen zurückgeführt werden.

Ein quantitativer Ansatz findet sich bei Fleming et al. (2006). Sie berücksichtigen die Heterogenität der Informationsrate und erweitern das modifizierte Modell von Andersen (1996), indem sie eine persistente sowie eine nicht persistente Komponente im Informationsfluss zulassen. Andersen (1996) modelliert die logarithmierte Informationsrate als einen AR(1)-Prozess (vgl. Gleichung 3.90), was von Fleming et al. (2006) als problematisch gesehen wird. Sollte die Informationsrate einem AR(1)-Prozess folgen, müsste sie das gleiche Ausmaß an Persistenz aufweisen wie lineare MMS (*minimum mean square*, mittlerer quadratischer Fehler) Prognosen der Varianzen von Renditen, und diese Prognosen sind identisch mit der bedingten Varianz aus dem GARCH(1,1)-Modell. GARCH-Modelle implizieren eine starke Persistenz mit der Halbwertszeit der Auswirkung der Schocks von geschätzten zwei bis sechs Monaten. Es wird zwar eine Persistenz in der Informationsrate erwartet, aber es ist nicht plausibel, dass sie so stark sein wird. Um die Verbindung zwischen

der Persistenz in der Prognose im Sinne des mittleren quadrierten Fehlers und der Persistenz in der Informationsrate zu unterbrechen, modellieren Fleming et al. (2006) den Informationsfluss als eine Summe aus einem AR(1)-Prozess und *white noise*. Das ermöglicht eine nicht persistente Komponente in der Varianz der Preisänderungen, die mit den verzögerten quadrierten Renditen nicht korreliert ist. Des weiteren wird noch ein Faktor in der Informationsrate modelliert, der als ein persistenter Fluss der öffentlichen Informationen interpretiert werden kann, der Preisänderungen ohne Handel hervorruft.

In Anlehnung an Andersen (1996) modellieren Fleming et al. (2006) zuerst die Preisänderungen und das Handelsvolumen als ein bivariates SV (*stochastic volatility*) Modell:

$$r_t = \sigma_t z_{rt}, \quad (4.1)$$

$$v_t = \tau(\mu + \lambda\sigma_t^2) + \tau(\mu + \lambda\sigma_t^2)^{1/2} z_{vt}. \quad (4.2)$$

Dabei bezeichnet z_{rt} *i.i.d.* $N(0,1)$ standardisierte Innovationen und z_{vt} solche *i.i.d.* standardisierten Innovationen, dass $v_t | \sigma_t^2 \sim \tau \cdot \text{Po}(\mu + \lambda\sigma_t^2)$. Außerdem ist z_{rt} und z_{vt} unabhängig für alle t und s . Ähnlich wie im Modell von Tauchen und Pitts (1983) ist im modifizierten Modell von Andersen (1996) die Varianz der Preisänderungen vom Informationsfluss untrennbar, was bedeutet, dass für die Dynamik der Informationsrate gleiche Restriktionen wie für die Varianz gelten.

Stochastische Volatilität und ARCH-Effekte

Zuerst betrachten Fleming et al. (2006) den Preisänderungsprozess (4.1). Unter der Annahme der Autokorrelation in der Informationsrate impliziert Gleichung (4.1), dass die Preisänderungen r_t mit einem SARV (*stochastic autoregressive volatility*) Modell beschrieben werden. Fleming et al. (2006) schlagen für die ökonometrische Analyse dieses Modells die Verwendung des Kalman-Filters vor, der zur Entfernung von Störungen aus den fehlerhaften Beobachtungen dient. Um die Benutzung dieser Methode zu ermöglichen, stellen Fleming et al. (2006) für r_t ein SARV-Modell mit einem Faktor auf:

$$r_t = \sigma_t z_{rt}, \quad (4.3)$$

$$\sigma_t^2 = \omega + \phi\sigma_{t-1}^2 + \gamma\eta_t. \quad (4.4)$$

Dabei ist $z_{rt} \sim \text{NID}(0,1)$, η_t ist *white noise* mit $\text{Var}(\eta_t) = 1 - \phi^2$ und η_t ist unabhängig von z_{rs} für alle t und s . Fleming et al. (2006) drücken die Varianz des Tages t wie folgt aus:

$$\sigma_t^2 = \zeta + \gamma b_t \quad (4.5)$$

$$\text{mit } \zeta = \frac{1}{1-\phi} \quad \text{und} \quad b_t = \phi b_{t-1} + \eta_t.$$

Somit wird die Dynamik von r_t^2 vom folgenden System gesteuert:

$$b_t = \phi b_{t-1} + \eta_t, \quad (4.6)$$

$$r_t^2 = \zeta + \gamma b_t + \varepsilon_t. \quad (4.7)$$

Dabei ist $\varepsilon_t = (\zeta + \gamma b_t)(z_{rt}^2 - 1)$ *white noise* mit $\text{Var}(\varepsilon_t) = 2(\zeta^2 + \gamma^2)$ und die Kovarianzen sind gleich Null:

$$\text{Cov}(\varepsilon_t, \eta_s) = \text{Cov}(\varepsilon_t, b_s) = 0 \quad \text{für alle } t \text{ und } s, \quad (4.8)$$

$$\text{Cov}(\eta_t, b_s) = \text{Cov}(\eta_t, r_s^2) = \text{Cov}(\varepsilon_t, r_s^2) \quad \text{für alle } t > s. \quad (4.9)$$

Das bedeutet, dass die Gleichungen (4.6) und (4.7) ein lineares Zustandsraummodell darstellen. Diese Reparametrisierung des SARV-Prozesses zu einem Modell mit einem dynamischen latenten Faktor erlaubt Schätzungen mit Verwendung des Kalman-Filters.

Der Kalman-Filter ist ein Satz von Gleichungen, der normalerweise benutzt wird, um MMS lineare Prognosen der Zustände eines Prozesses basierend an vergangenen Beobachtungen zu machen. Für das System (4.6) und (4.7) ergibt der Filter MMS lineare Prognosen von b_{t+1} gegeben $\{r_t^2, r_{t-1}^2, \dots, r_1^2\}$ für $t = 1, 2, \dots, T$, was auch MMS lineare Prognosen für σ_{t+1}^2 bei

$t = 1, 2, \dots, T$ impliziert. Somit ergibt sich die Prognose für σ_{t+1}^2 des Tages t , die als $\sigma_{t+1|t}^2$ bezeichnet wird, wie folgt:

$$\sigma_{t+1|t}^2 = \zeta + \phi(\sigma_{t|t-1}^2 - \zeta) + \gamma K_t (r_t^2 - \sigma_{t|t-1}^2), \quad (4.10)$$

wobei K_t die Korrektur aus dem Kalman-Filter darstellt. Diese Prognose kann auch folgendermaßen dargestellt werden:

$$\sigma_{t+1|t}^2 = \omega + \beta_t \sigma_{t|t-1}^2 + \alpha_t r_t^2 \quad (4.11)$$

mit $\beta_t = (\phi - \gamma K_t)$ und $\alpha_t = \gamma K_t$.

Die Koeffizienten in der Gleichung (4.10) stellen bei der Initialisierung des Filters zu $t = 0$ Variablen dar. Mit zunehmendem t konvergieren diese Koeffizienten schnell zu Konstanten. Im stabilen Zustand haben die Varianzprognosen mit den Kalman-Filter gleiche Struktur wie bedingte Varianzen aus dem GARCH(1,1)-Modell.

Durch die Anwendung des linearen Filters decken Fleming et al. (2006) das Verhältnis zwischen der SARV-Dynamik, Persistenz in der Informationsrate und ARCH-Effekten in den Preisänderungen auf. Besonders wichtig dabei ist, dass MMS lineare Prognosen des σ_t^2 die gleiche Persistenz wie der Informationsfluss aufweisen, und wenn σ_t^2 einem AR(1)-Prozess folgt, nehmen diese Prognosen die gleiche Form wie bedingte Varianzen aus dem GARCH-Modell an.

Diese Verbindung erscheint problematisch, da Parameterschätzungen aus GARCH-Modellen eine starke Persistenz mit der Halbwertszeit der Auswirkung der Schocks von zwei bis sechs Monaten implizieren. Da MMS lineare Prognosen von σ_t^2 einen ähnlichen Grad an Persistenz aufweisen, bedeutet das unter der Annahme des AR(1)-Prozesses, dass die Halbwertszeit der Auswirkung der Informationsschocks auch von zwei bis sechs Monaten beträgt. Es wird zwar eine Persistenz in der Informationsrate erwartet, aber es ist nicht plausibel, dass sie so stark ist. Um die Verbindung zwischen der Persistenz in der MMS linearen Prognose und der Persistenz in der Informationsrate zu unterbrechen, wenden sich Fleming et al. (2006) einem

SARV-Modell mit mehr als einem Faktor zu und spezifizieren den Informationsfluss als eine Summe aus einem AR(1)-Prozess und *white noise*:

$$\sigma_t^2 = h_t + \rho c_t. \quad (4.12)$$

Dabei ist c_t eine *white noise* Innovation mit $\text{Var}(c_t)=1$ und $h_t = E(\sigma_t^2 | \mathfrak{I}_{t-1})$ die bedingte Varianz mit der Informationsmenge \mathfrak{I}_t des Tages t . Diese Spezifikation ermöglicht die Berücksichtigung einer nicht persistenten Komponente in der Varianz der Preisänderungen, die mit den verzögerten quadrierten Renditen nicht korreliert ist. Fleming et al. (2006) nehmen wiederum an, dass b_t einem AR(1)-Prozess folgt, und drücken die bedingte Varianz wie $h_t = \zeta + \gamma b_t$ aus, um folgendes lineares Zustandsraummodell für r_t^2 zu erhalten:

$$b_t = \phi b_{t-1} + \eta_{bt}, \quad (4.13)$$

$$c_t = \eta_{ct}, \quad (4.14)$$

$$r_t^2 = \zeta + \gamma b_t + \rho c_t + \varepsilon_t. \quad (4.15)$$

Dabei ist $\text{Var}(\eta_{bt})=1-\phi^2$, $\text{Var}(\eta_{ct})=1$ und $\text{Var}(\varepsilon_t)=2(\zeta^2 + \gamma^2 + \rho^2)$. Diese Spezifikation mit zwei Faktoren impliziert, dass σ_t^2 eine ARMA(1,1)-Dynamik besitzt. Das bedeutet, dass dieses Modell eine starke Persistenz in der bedingten Varianz nachbilden kann, ohne eine starke Persistenz in der Informationsrate nach sich zu ziehen.

Fleming et al. (2006) zeigen den Vorteil dieses Zwei-Faktoren-Modells bei den gefilterten Schätzern der Varianz σ_t^2 gegeben $\{r_t^2, r_{t-1}^2, \dots, r_1^2\}$. Das Modell mit einem Faktor impliziert folgende MMS lineare Schätzung von σ_t^2 am Tag t :

$$\sigma_{t|t}^2 = \sigma_{t|t-1}^2 + \gamma \kappa_t e_t. \quad (4.16)$$

Dabei ist $e_t = r_t^2 - \sigma_{t|t-1}^2$ der Fehler in der Prognose von r_t^2 und κ_t ein Koeffizient, der proportional zur Korrektur aus dem Kalman-Filter ist: $K_t = \phi\kappa_t$. Das Modell mit zwei Faktoren impliziert Folgendes:

$$\sigma_{t|t}^2 = \sigma_{t|t-1}^2 + \gamma\kappa_{bt}e_t + \rho\kappa_{ct}e_t. \quad (4.17)$$

Dabei sind κ_{bt} und κ_{ct} solche Koeffizienten, dass das erste Element des Korrekturvektors mit $K_{bt} = \phi\kappa_t$ gegeben ist. Der zweite Koeffizient κ_{ct} fließt in den letzten Summanden der Gleichung (4.17) ein. Dieser Term bildet die nicht persistente Komponente von σ_t^2 ab, da er in der Gleichung (4.10) für $\sigma_{t+1|t}^2$ nicht vorhanden ist.

Stochastische Volatilität, Handelsvolumen und ARCH-Effekte

Im nächsten Schritt berücksichtigen Fleming et al. (2006) die Gleichung (4.2) des Handelsvolumens. Innovationen im Handelsvolumen sind mit $\varepsilon_{vt} = \tau(\mu + \lambda\sigma_t^2)^{1/2} z_{vt}$ gegeben. Die Dynamik von r_t^2 und v_t ist im Zwei-Faktoren-Modell wie folgt spezifiziert:

$$b_t = \phi b_{t-1} + \eta_{bt} \quad (4.18)$$

$$c_t = \eta_{ct}, \quad (4.19)$$

$$r_t^2 = \zeta_r + \gamma_r b_t + \rho_r c_t + \varepsilon_{rt}, \quad (4.20)$$

$$v_t = \zeta_v + \gamma_v b_t + \rho_v c_t + \varepsilon_{vt} \quad (4.21)$$

$$\text{mit } \zeta_v = \tau(\mu + \lambda\zeta_r), \quad \gamma_v = \tau\lambda\gamma_r, \quad \rho_v = \tau\lambda\rho_r, \\ \text{Var}(\eta_{ct}) = 1, \quad \text{Var}(\eta_{ct}) = 2(\zeta_r^2 + \gamma_r^2 + \rho_r^2), \quad \text{Var}(\varepsilon_{vt}) = \tau^2(\mu + \lambda\zeta_r).$$

Die *white noise* Komponente ε_{vt} besitzt folgende Kovarianzen:

$$\text{Cov}(\varepsilon_{vt}, \eta_{bs}) = \text{Cov}(\varepsilon_{vt}, b_s) = \text{Cov}(\varepsilon_{vt}, \eta_{cs}) = \text{Cov}(\varepsilon_{vt}, c_s) = 0 \quad \text{für alle } t \text{ und } s, \quad (4.22)$$

$$\text{Cov}(\varepsilon_{vt}, r_s^2) = \text{Cov}(\varepsilon_{vt}, v_s) = 0 \quad \text{für alle } t > s. \quad (4.23)$$

Somit stellt das System (4.18)-(4.21) ein lineares Zustandsraummodell dar. Der Faktor b_t folgt dabei einem AR(1)-Prozess und erfasst die gemeinsame persistente Komponente in r_t^2 und v_t , während der Faktor c_t die gemeinsame nicht persistente Komponente in diesen Variablen erfasst. Durch die Aufnahme der Parameter ρ_r und ρ_v in das Modell wird es möglich, jene Komponente von r_t^2 und v_t zu isolieren, die den nicht persistenten Teil der Informationsrate widerspiegelt.

Verallgemeinerung des bivariaten Modells

Um das Modell weiter zu verallgemeinern, führen Fleming et al. (2006) einen dritten Faktor ein, der als eine zusätzliche Komponente der Informationsrate interpretiert wird, und erhalten folgende Spezifikation:

$$a_t = \phi_a a_{t-1} + \eta_{at} \quad (4.24)$$

$$b_t = \phi_b b_{t-1} + \eta_{bt} \quad (4.25)$$

$$c_t = \eta_{ct}, \quad (4.26)$$

$$r_t^2 = \zeta_r + \delta a_t + \gamma_r b_t + \rho_r c_t + \varepsilon_{rt}, \quad (4.27)$$

$$v_t = \zeta_v + \gamma_v b_t + \rho_v c_t + \varepsilon_{vt}. \quad (4.28)$$

Dabei sind η_{at} , η_{bt} und η_{ct} von einander unabhängige *white noise* Innovationen. Durch die Aufnahme des Faktors a_t in das Modell wird die Erfassung einer persistenten Komponente der Volatilität ermöglicht, die in keiner linearen Beziehung mit der persistenten Komponente des Handelsvolumens steht.

Fleming et al. (2006) bieten zwei Interpretationen für diese Spezifikation an. Erstens, kann der dritte Faktor a_t als eine ad hoc Erweiterung des Zwei-Faktoren-Modells gesehen werden. In diesem Fall wäre ein statistisch signifikanter Schätzer des entsprechenden Parameters δ ein Beweis gegen das modifizierte Mischungsverteilungsmodell. Die zweite Erklärung ist für diese Arbeit von besonderem Interesse. Der Faktor a_t wird als eine zusätzliche Komponente der Informationsrate interpretiert, nämlich ein Informationsfluss der öffentlichen Informationen, der Preisänderungen ohne Handel verursacht. Aus dieser Sicht ist die Spezifikation konsistent mit dem modifizierten Mischungsverteilungsmodell, welches das Eintreffen öffentlicher sowie privater Informationen erfasst. Unabhängig davon, ob diese Spezifikation mit der modifizierten MDH konsistent ist oder nicht, erlaubt sie neue Erkenntnisse über die Beziehung zwischen den ARCH-Effekten und dem Handelsvolumen. Wenn der dritte Faktor a_t die Anpassung des Modells an die Daten verbessert, bedeutet das, dass quadrierte Preisänderungen solche persistente Komponente in der Volatilität erfassen, die vom Handelsvolumen nicht erfasst wird.

Fleming et al. (2006) testen verschiedene Spezifikationen des modifizierten Mischungsverteilungsmodells mit der *quasi maximum likelihood* (QML) Methode. Der zu schätzende Parametervektor ist gegeben mit $\theta = (\phi, \zeta, \gamma, \rho)'$. Die Quasi Loglikelihoodfunktion lautet

$$L(\theta) = \sum_{t=1}^T l_t(\theta) \quad (4.29)$$

$$\text{mit } l_t = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log \xi_{t|t-1}^2 - \frac{1}{2} \left(\frac{r_t^2 - \sigma_{t|t-1}^2}{\xi_{t|t-1}} \right)^2$$

$$\text{und } \xi_{t|t-1}^2 = \gamma^2 P_{b,t|t-1} + \rho^2 + 2(\zeta^2 + \gamma^2 + \rho^2).$$

Dabei ist $P_{b,t|t-1}$ die mittlere quadratische Abweichung von $b_{t|t-1}$. Die bedingte Verteilung von r_t^2 ist $N(\sigma_{t|t-1}^2, \xi_{t|t-1}^2)$, wenn η_{bt} , η_{ct} und ε_t multivariat normalverteilt sind. Grundsätzlich

ergibt sich der QML-Schätzer aus der Minimierung der mittleren quadratischen Abweichung in der Prognose für r_t^2 .

Um ein Benchmark zu erhalten, schätzen Fleming et al. (2006) zuerst ein GARCH(1,1)-Modell in der Form

$$r_t = \sqrt{h_t} z_t, \quad (4.30)$$

$$h_t = \omega + \beta h_{t-1} + \alpha r_{t-1}^2. \quad (4.31)$$

Fleming et al. (2006) untersuchen Tagesrenditen und Handelsvolumen von 20 Aktien des MMI⁷ für die Periode zwischen 1.01.1988 und 31.11.2000. Die Schätzergebnisse zeigen Werte für α unter 0,1 und für β über 0,8, was auf die ARCH-Effekte hindeutet. Die bedingte Varianz h_t ist stark persistent mit Werten über 0,9 für die meisten Aktien. Fleming et al. (2006) finden, dass das GARCH(1,1)-Modell eine geringe Erklärungskraft besitzt, da das Verhältnis $\text{Var}(h_t)/\text{Var}(r_t^2)$ zwischen 1% und 6% liegt. Die Anpassungsgüte des Modells wird mit der Ljung-Box Statistik unter der Nullhypothese „keine Autokorrelation“ getestet. Die Statistik $r_t^2 - h_t$ ist in zwölf aus zwanzig Fällen zum Niveau von 5% und in neun Fällen zum Niveau von 1% signifikant. Somit wird die Nullhypothese in den meisten Fällen verworfen. Die Ljung-Box Statistik für die Autokorrelation in r_t^2/h_t ist in fünf Fällen signifikant zum Niveau von 5%. Wenn die auf dem modifizierten Mischungsverteilungsmodell basierenden SARV-Modelle gut spezifiziert sind, sind die Residuen der MMS Prognosen von r_t^2 *white noise*. Somit ist zu erwarten, dass die Autokorrelationen in den SARV Prognosefehlern weniger stark als beim Benchmark Modell sind.

Nach dem GARCH(1,1)-Modell wenden sich Fleming et al. (2006) der ersten SARV-Spezifikation zu – dem Modell mit zwei Faktoren, in welchem die Varianz der Preisänderungen (bzw. der Informationsfluss) ARMA(1,1)-Dynamik aufweist. Dieses Modell

⁷ Fleming et al. (2006) haben folgende Aktien untersucht: American Express (AXP), AT&T (T), Chevron (CHV), Coca-Cola (KO), Disney (DIS), Dow Chemical (DOW), DuPont (DD), Eastman Kodak (EK), Exxon-Mobil (XOM), General Electric (GE), General Motors (GM), International Business Machines (IBM), International Paper (IP), Johnson & Johnson (JNJ), McDonald's (MCD), Merck (MRK), 3M (MMM), Philip Morris (MO), Procter and Gamble (PG), und Sears (S).

wird unter Verwendung des Zustandsraummodells (4.13)-(4.15) angepasst. Der geschätzte Wert des Faktors ϕ liegt über 0,9 für die meisten Aktien und weist auf eine starke Persistenz in b_t hin. Dieser Wert entspricht ungefähr der Summe aus α und β aus dem GARCH(1,1)-Modell. Das Zwei-Faktoren-Modell impliziert, dass die MMS linearen Prognosen von σ_t^2 der bedingten Varianz h_t entsprechen. Der Persistenzfaktor b_t erfasst in der Tat ARCH-Effekte in täglichen Preisänderungen.

Gleichzeitig erfasst der nicht persistente Faktor c_t eine andere Komponente in der Varianz der Preisänderungen. Der Parameter ρ misst den Beitrag des nicht persistenten Faktors zu σ_t^2 . Der geschätzte Wert von ρ ist signifikant zum Niveau von 5% für 18 der 20 Aktien. Für 13 Aktien ist ρ sogar größer als γ . Wenn man bedenkt, dass die unbedingte Varianz von σ_t^2 mit $\gamma^2 + \rho^2$ gegeben ist, bedeutet das, dass c_t mehr als die Hälfte der Varianz von σ_t^2 erfasst. Durch die Aufnahme des Faktors c_t ins Modell wird somit eine große nicht persistente Komponente der Varianz der Preisänderungen erfasst, die weder mit der Ein-Faktor-SARV-Spezifikation noch mit dem GARCH-Modell erfasst wird.

Die Ljung-Box Statistik zeigt, dass die Zwei-Faktoren-SARV-Spezifikation eine zumindest gleich gute Anpassung ermöglicht, wie das GARCH(1,1)-Modell: Die Statistik für $r_t^2 - h_t$ ist nur für neun Aktien zum Niveau von 5% signifikant (im GARCH(1,1)-Modell waren das zwölf). Die Schätzer für $\text{Var}(\sigma_t^2)/\text{Var}(r_t^2)$ sind mit 13% bis 30% höher als im GARCH(1,1)-Modell.

Nachdem die Schätzung des Zwei-Faktoren-Modells die Existenz der zwei verschiedenen Komponenten in der Varianz der Preisänderungen bestätigt hat, nehmen Fleming et al. (2006) das Handelsvolumen in die Spezifikation auf, um zu testen, ob diese Ergebnisse mit der modifizierten MDH konsistent sind.

Der geschätzte Wert für ϕ ist kleiner, als in der univariaten SARV-Spezifikation. Das ist das erste Anzeichen, dass das modifizierte Mischungsverteilungsmodell für die Daten nicht passt. Die Aufnahme des Handelsvolumens ändert nur wenig an den Schätzungen für γ und ρ im Vergleich zur univariaten Spezifikation. Die Schätzer für $\text{Var}(\sigma_t^2)/\text{Var}(r_t^2)$ ändern sich auch

nicht wesentlich, was bedeutet, dass die Aufnahme des Handelsvolumens keine Auswirkung auf die Ergebnisse bezüglich der Persistenz in der Varianz hat. Dafür ist eine leichte Verbesserung in der Prognose und eine wesentliche Verbesserung im Filtern zu verzeichnen. Das bedeutet, dass das Handelsvolumen aufschlussreich über die nicht persistente Komponente der Volatilität ist. Um das zu untersuchen, vergleichen Fleming et al. (2006) diverse Bestimmtheitsmaße und kommen zum Schluss, dass die nicht prognostizierbare Komponente des Handelsvolumens eine wesentlich bessere Proxyvariable für die nicht persistente Komponente der Varianz ist, als die nicht prognostizierbare Komponente der quadrierten Renditen, die fast keine zusätzliche Erklärungskraft besitzt.

Obwohl diese Ergebnisse durchaus interessant sind, passt dieses Model nicht zu den Daten. Die Ljung-Box Statistik für quadrierte Renditen ist signifikant für 17 Aktien zum Niveau von 5%. Das bedeutet, dass sich die Qualität der r_t^2 Prognose mit der Aufnahme des Handelsvolumens verschlechtert. Die Statistik für die serielle Korrelation in den Residuen der v_t Prognose ist signifikant für 18 Aktien zum gleichen Signifikanzniveau.

Um die scheinbare Unfähigkeit der modifizierten MDH, die gemeinsame Dynamik des Handelsvolumens und der Volatilität zu erfassen, weiter zu untersuchen, schätzen Fleming et al. (2006) eine weniger restriktive bivariate SARV-Spezifikation mit drei Faktoren. Diese Spezifikation schließt das Zwei-Faktoren-Modell als einen Spezialfall ein. Der dritte Faktor a_t beeinflusst nur die quadrierten Preisänderungen und erlaubt somit eine Persistenz in der Volatilität, die in keinem Zusammenhang mit der Persistenz im Handelsvolumen steht.

Die geschätzten Werte der Parameter sind den Werten aus dem Zwei-Faktoren-Modell mit dem Handelsvolumen sehr ähnlich. Der Schätzer des Parameters ϕ_b , der die Autokorrelation im Faktor b_t misst, der gemeinsame Persistenz in r_t^2 und v_t erfasst, bleibt fast unverändert. Die Koeffizienten γ_r und γ_v des persistenten Faktors b_t und die Koeffizienten ρ_r und ρ_v des nicht persistenten Faktors c_t bleiben auch im Wesentlichen gleich. Damit erklären Fleming et al. (2006) die Verschlechterung der r_t^2 Prognose nach der Aufnahme des Handelsvolumens in das Zwei-Faktoren-Modell. Die Prognosefehler der quadrierten Renditen haben eine vergleichsweise geringe Gewichtung bekommen, weil die Prognosefehler des Handelsvolumens wesentlich genauere Signale über die Volatilität enthalten. Deswegen hat

die Einführung des neuen Faktors a_t , der nur für die quadrierten Preisänderungen relevant ist, wenig Auswirkungen auf die Parameter aus dem Zwei-Faktoren-Modell.

Die Schätzungen für den dritten Faktor a_t zeigen, dass das Drei-Faktoren-Modell eine wichtige Komponente der Volatilität erfasst, die im Zwei-Faktoren-Modell fehlt. Die geschätzten Werte für den Faktor a_t sind immer signifikant. Dieser Faktor ist persistenter als der gemeinsame Faktor für die quadrierten Preisänderungen und Handelsvolumen. Die geschätzte Autokorrelation in a_t beträgt mindestens 0,98 für jede Aktie. Somit zeigt das Modell, dass es eine stark persistente Komponente in der Volatilität der Preisänderungen gibt, die durch die gemeinsame Persistenz in den quadrierten Preisänderungen und Handelsvolumina nicht erklärt werden kann.

Fleming et al. (2006) finden, dass die Güte der Volatilitätsprognose in diesem Modell ähnlich wie im GARCH(1,1)-Modell ist: Die Ljung-Box Statistik für quadrierte Preisänderungen ist in fünf Fällen signifikant zum Niveau von 5%. Die Ergebnisse der Handelsvolumenprognose sind schlechter: Die Ljung-Box Statistik für Residuen ist in 18 Fällen signifikant zum Niveau von 5%. Das bedeutet, dass die Struktur der Autokorrelation des Handelsvolumens komplexer ist, als das ARMA(1,1)-Modell impliziert.

Zusammenfassend lässt sich feststellen, dass die Ergebnisse von Fleming et al. (2006) die Heterogenität der Informationsrate bestätigen.

4.3 Zeitabhängige Informationssensibilität

Liesenfeld (2001) lockert die Annahmen des Modells von Tauchen und Pitts (1983) (vgl. Abschnitt 3.4.1), indem er eine zeitabhängige Informationssensibilität der Marktteilnehmer zulässt. In seinem erweiterten Modell sind die latente Informationsrate sowie die latente Sensibilität autokorrelierte Zufallsvariablen. Um die Darstellung übersichtlich zu machen, werden hier die relevanten Gleichungen des Modells aus dem Abschnitt 3.4.1 nochmals angeführt. Zu beachten ist, dass Liesenfeld (2001) im Gegenteil zu Tauchen und Pitts (1983) mit logarithmierten Preisen arbeitet.

Die Veränderung des Reservationspreises $\Delta P_{ij}^* = P_{ij}^* - P_{i-1,j}^*$ wird in zwei Komponenten zerlegt:

$$\Delta P_{ij}^* = \phi_i + \psi_{ij}, \quad (4.32)$$

$$\text{mit } E[\phi_i] = E[\psi_{ij}] = 0, \quad \text{Var}[\phi_i] \equiv \sigma_\phi^2, \quad \text{Var}[\psi_{ij}] \equiv \sigma_\psi^2.$$

Dabei sind ϕ und ψ zeit- und händlerunabhängig. Die Komponente ϕ_i ist gleich für alle Händler, die Komponente ψ_{ij} ist spezifisch für den j -ten Händler. Die Variablen σ_ϕ^2 und σ_ψ^2 messen die Sensibilität, mit welcher die Reservationspreise auf die Ankunft der preisrelevanten Informationen reagieren. Tauchen und Pitts (1983) spezifizieren die ersten zwei Momente für die Preisänderungen r_i und das Handelsvolumen v_i :

$$\mu_r \equiv E[r_i] = 0, \quad (4.33)$$

$$\sigma_r^2 \equiv \text{Var}[r_i] = \sigma_\phi^2 + \frac{\sigma_\psi^2}{J}, \quad (4.34)$$

$$\mu_v \equiv E[v_i] = \left(\frac{\alpha}{2}\right) \sigma_\psi \sqrt{\frac{2}{\pi}} \left(\sqrt{\frac{J-1}{J}}\right) J, \quad (4.35)$$

$$\sigma_v^2 \equiv \text{Var}[v_i] = \left(\frac{\alpha}{2}\right)^2 \sigma_\psi^2 \left(1 - \frac{2}{\pi}\right) J. \quad (4.36)$$

Preisänderungen und Handelsvolumen sind in diesem Modell seriell unabhängig. Aus den Gleichungen (4.34)-(4.36) wird ersichtlich, dass die Momente von r_i und v_i Funktionen der Varianz des händlerspezifischen Schocks σ_ψ^2 sind. Die Varianz des für alle gleichen Schocks σ_ϕ^2 fließt dagegen nur in die Varianzgleichung der Preisänderungen (4.34) ein. Der Grund dafür ist, dass die gemeinsame Komponente ϕ_i keinen Einfluss auf das Handelsvolumen hat und nur für den Preisänderungsprozess relevant ist.

Angenommen die Anzahl der Händler J und die Variablen σ_ϕ^2 und σ_ψ^2 bleiben konstant über die Zeit, dann lässt sich die Verteilung der Preisänderungen und des Handelsvolumens als eine bivariate Normalverteilung mit der Informationsrate als Mischungsvariable folgendermaßen darstellen:

$$r_t | n_t \sim N(0, \sigma_r^2 n_t) \quad (4.37)$$

$$v_t | n_t \sim N(\mu_0 + \mu_v n_t, \sigma_v^2 n_t) \quad (4.38)$$

mit $\text{Cov}(r_t, v_t | n_t) = 0$

Liesenfeld (2001) führt die zusätzliche Komponente μ_0 in das Modell ein, um den von der Ankunft der Informationen unabhängigen Teil des Handelsvolumens (vgl. Abschnitt 3.4.2) nachzubilden. Die bedingte Varianz der Preisänderungen wird durch die Mischungsvariable bestimmt. Das bedeutet, dass das Zeitreihenverhalten der Volatilität der Preisänderungen vom Zeitreihenverhalten der Informationsrate abhängt, welches auch die Dynamik des Handelsvolumens beeinflusst. Daraus folgt, dass die im Varianzprozess beobachtbare Persistenz ein Resultat der Persistenz im Informationsprozess sein kann. Die Persistenz im Informationsprozess ist leicht erklärbar: Auf eine unerwartete Ankündigung folgen oft weitere, die mit der ersten thematisch verbunden sind. Liesenfeld (2001) modelliert die Autokorrelation im Informationsprozess als einen AR(1)-Prozess für den Logarithmus der Informationsrate $\lambda_t = \ln n_t$:

$$\lambda_t | \lambda_{t-1} \sim N(\delta_\lambda \lambda_{t-1}, v_\lambda^2). \quad (4.39)$$

Dabei ist δ_λ der Persistenzparameter der Volatilität. Das bivariate Modell (4.37)-(4.39) impliziert eine univariate Spezifikation der Preisänderungen wie im ARV-Modell (vgl. Abschnitt 3.3.1). Untersuchungen von Andersen (1996) und Liesenfeld (1998) zeigen, dass der geschätzte Persistenzparameter δ_λ im bivariaten Modell (4.37)-(4.39) signifikant kleiner ist als in der univariaten Spezifikation der Preisänderungen (4.37), (4.39). Ein möglicher Grund dafür ist, dass Preisänderungen und Handelsvolumen einen unterschiedlichen Grad an Persistenz aufweisen, der mit einem einfachen Mischungsverteilungsmodell mit einer latenten Variablen wie Informationsrate nicht nachzubilden ist.

Die zeitabhängige Spezifikation der Varianzen σ_ϕ^2 und σ_ψ^2 impliziert, dass die Informationssensibilität der Reservationspreise der Händler über die Zeit nicht konstant ist. Zum Beispiel, die Unsicherheit über die jetzige und künftige Wirtschaftslage wird oft als ein Faktor betrachtet, der die Volatilität der Preisänderungen der Aktien beeinflusst. Liesenfeld (2001) nimmt an, dass die zeitabhängigen Varianzen $\sigma_{\phi,t}^2$ und $\sigma_{\psi,t}^2$ gemeinsam von einem latenten Prozess ω_t gesteuert werden, der den Grad der Unsicherheit, der diese Varianzen beeinflusst, widerspiegelt. Liesenfeld (2001) spezifiziert die zeitabhängigen Varianzen folgendermaßen, um deren Nichtnegativität sicherzustellen:

$$\ln(\sigma_{\phi,t}^2) = \gamma_\phi + \alpha_\phi \omega_t, \quad (4.40)$$

$$\ln(\sigma_{\psi,t}^2) = \gamma_\psi + \alpha_\psi \omega_t. \quad (4.41)$$

Nach dem Einfügen der Varianzgleichungen (4.40) und (4.41) in die Gleichungen der Momente der Preisänderungen und des Handelsvolumens (4.34)-(4.36), ergibt sich folgende Spezifikation des Mischungsverteilungsmodells (4.37) und (4.38):

$$r_t \mid \lambda_t, \omega_t \sim N\left(0, \left[\beta_1 e^{\alpha_\phi \omega_t} + \beta_2 e^{\alpha_\psi \omega_t}\right] e^{\lambda_t}\right), \quad (4.42)$$

$$v_t \mid \lambda_t, \omega_t \sim N\left(\mu_0 + \left[\beta_3 e^{\alpha_\psi \omega_t / 2}\right] e^{\lambda_t}, \left[\beta_4 e^{\alpha_\psi \omega_t}\right] e^{\lambda_t}\right). \quad (4.43)$$

Dabei beeinflusst die Sensibilität, mit welcher auf die individuellen Schocks $\sigma_{\psi,t}^2$ bzw. $e^{\alpha_\psi \omega_t}$ reagiert wird, die Preisänderungen sowie das Handelsvolumen. Die Sensibilität für die gemeinsamen Schocks $\sigma_{\phi,t}^2$ bzw. $e^{\alpha_\phi \omega_t}$ geht dagegen nur in die Preisänderungsgleichung ein. Die Variablen β_1 , β_2 , β_3 und β_4 sind positive Parameter. Aus den Gleichungen (4.42) und (4.43) folgt, dass das gemeinsame Verhalten des Handelsvolumens und der Volatilität der Preisänderungen gleichzeitig von der Informationsrate e^{λ_t} und vom Grad der Unsicherheit ω_t abhängt. Um den Prozess ω_t zu spezifizieren, berücksichtigt Liesenfeld (2001) Folgendes. Der Prozess, der die Ankunft neuer Informationen steuert, kann die hohe Persistenz in der Volatilität nicht vollständig erklären. Das deutet auf die Existenz anderer autokorrelierten Faktoren hin. Außerdem muss der von Nelson (1991) beschriebene Leverage-Effekt, wonach

der Rückgang im Preis typischerweise von einem höheren Anstieg der Volatilität gefolgt wird, als eine entsprechende Erhöhung des Preises, Berücksichtigung finden. Um das zu gewährleisten, lässt Liesenfeld (2001) eine Abhängigkeit des Prozesses ω_t von den verzögerten Preisänderungen r_{t-1} zu. Diese Asymmetrie in der Reaktion der Preise kann als Informationssensibilität interpretiert werden: Händler reagieren stärker auf „schlechte Nachrichten“ (negative Preisänderungen) als auf „gute Nachrichten“ (positive Preisänderungen). Basierend auf diesen Überlegungen, spezifiziert Liesenfeld (2001) den latenten Prozess ω_t wie folgt:

$$\omega_t | \omega_{t-1}, r_{t-1} \sim N(\kappa r_{t-1} + \delta_\omega \omega_{t-1}, v_\omega^2). \quad (4.44)$$

Es wird angenommen, dass ω_t und λ_t gegebene verzögerte Preisänderungen unkorreliert sind. Um die Spezifikation zu vervollständigen, führt Liesenfeld (2001) eine zusätzliche Restriktion $\alpha_\phi \equiv 1$ in die Gleichungen (4.40)-(4.42) ein. Das impliziert, dass $\sigma_{\phi,t}^2$ dem Ausdruck e^{ω_t} und $\sigma_{\psi,t}^2$ dem Ausdruck $e^{\tilde{\alpha}_\psi \omega_t}$ entspricht. Dann sind die zu schätzenden Parameter $\tilde{\alpha}_\psi = \alpha_\psi / \alpha_\phi$, $\tilde{\kappa} = \alpha_\phi \kappa$ und $\tilde{v}_\omega = \alpha_\phi v_\omega$. Schlussendlich nimmt das erweiterte Mischungsverteilungsmodell von Liesenfeld (2001) folgende Gestalt an:

$$r_t | \lambda_t, \omega_t \sim N(0, [\beta_1 e^{\omega_t} + \beta_2 e^{\tilde{\alpha}_\psi \omega_t}] e^{\lambda_t}), \quad (4.45)$$

$$v_t | \lambda_t, \omega_t \sim N(\mu_0 + [\beta_3 e^{\tilde{\alpha}_\psi \omega_t / 2}] e^{\lambda_t}, [\beta_4 e^{\tilde{\alpha}_\psi \omega_t}] e^{\lambda_t}), \quad (4.46)$$

mit

$$\omega_t | \omega_{t-1}, r_{t-1} \sim N(\tilde{\kappa} r_{t-1} + \delta_\omega \omega_{t-1}, \tilde{v}_\omega^2), \quad (4.47)$$

$$\lambda_t | \lambda_{t-1} \sim N(\delta_\lambda \lambda_{t-1}, v_\lambda^2). \quad (4.48)$$

Im Gegenteil zum Modell von Tauchen und Pitts (1983) erlaubt dieses erweiterte Modell eine heterogene Dynamik der Preisänderungen und des Handelsvolumens durch die Einführung des zusätzlichen Prozesses ω_t . Dieser Prozess hat unterschiedliche Auswirkungen auf die

Preisänderungen und das Handelsvolumen: Er beeinflusst die Volatilitätsdynamik durch die Sensibilität der Händler gegenüber den gemeinsamen sowie spezifischen Schocks und das Handelsvolumen nur durch die Sensibilität gegenüber den spezifischen Schocks. Der Parameter $\tilde{\alpha}_\psi$ misst die relative Wichtigkeit des zeitabhängigen Faktors ω_t für die Sensibilität gegenüber den beiden Arten von Schocks. Zum Beispiel, wenn $\tilde{\alpha}_\psi = 0$ und $\tilde{v}_\omega > 0$, dann ist die Sensibilität auf spezifische Schocks $\sigma_{\psi,t}^2$ über die Zeit konstant und die Sensibilität auf gemeinsame Schocks $\sigma_{\phi,t}^2$ zeitabhängig. In diesem Fall würde ω_t nur den Volatilitätsprozess und nicht das Handelsvolumen beeinflussen. Wenn $\tilde{\kappa} = \delta_\omega = \tilde{v}_\omega = 0$, dann sind $\sigma_{\psi,t}^2$ und $\sigma_{\phi,t}^2$ konstant über die Zeit und das Modell entspricht dem Ausgangsmodell von Tauchen und Pitts (1983) mit einem latenten zeitabhängigen Faktor. Parameter $\tilde{\kappa}$ und $\tilde{\alpha}_\psi \tilde{\kappa}$ messen die Asymmetrie der Sensibilität gegenüber den gemeinsamen und spezifischen Schocks respektive. Bei negativen Werten von $\tilde{\kappa}$ und $\tilde{\alpha}_\psi \tilde{\kappa}$ führen negative Preisänderungen zu höheren Informationssensibilität als positive Preisänderungen. Positive Werte von β_1 , β_2 und β_3 implizieren dann bei fallenden Preisen höhere Volatilität und Handelsvolumen als bei steigenden Preisen. Aus den Gleichungen (4.47) und (4.48) ist ersichtlich, dass bei $\delta_\omega < \delta_\lambda$ der Faktor λ_t die *long-run* Komponente und ω_t die *short-run* Komponente der Volatilität repräsentiert, und umgekehrt.

Die latenten Variablen λ_t und ω_t sind im Ausgangsmodell (4.37) und (4.38) und im erweiterten Modell (4.45)-(4.48) autokorreliert, deswegen lassen sich diese Modelle nicht mit der einfachen ML-Methode schätzen. Liesenfeld (2001) greift auf die SML-Methode zurück. Dafür wird zuerst die Matrix der beobachtbaren Variablen aus der gemeinsamen Verteilung rausintegriert:

$$L(\theta; \mathbf{Y}_T) = \int f(\mathbf{Y}_T, \mathbf{X}_T | \theta) d\mathbf{X}_T . \quad (4.49)$$

Dabei ist $\mathbf{Y}_T = \{y_t\}_{t=1}^T$ die Matrix der beobachtbaren Variablen mit $y_t = (r_t, v_t)$;

$\mathbf{X}_t = \{x_t\}_{t=1}^T$ ist die Matrix der latenten Variablen mit $x_t = \lambda_t$ für das Ausgangsmodell und $x_t = (\lambda_t, \omega_t)$ für das erweiterte Modell;

θ ist der zu schätzende Parametervektor.

Diese Likelihoodfunktion ist ein hochdimensionaler Integral mit T Dimensionen für das Ausgangsmodell und $2T$ Dimensionen für das erweiterte Modell.

Liesenfeld (2001) schätzt die Likelihoodfunktion (4.49) mit der MC Methode und maximiert sie anschließend für θ .

Um die Ergebnisse der Schätzungen der bivariaten Modelle vergleichen zu können, schätzt Liesenfeld (2001) zuerst das an seine Daten angepasste ARV-Modell und GARCH(1,1)-Modell mit SML- und ML-Methode respektive.

Die Schätzung des Ausgangsmodells (4.37) und (4.38) zeigt signifikante Werte für alle Parameter, einschließlich jene des Informationsprozesses δ_λ und ν_λ . Die Schätzungen für δ_λ sind aber wesentlich kleiner als im univariaten ARV-Modell. Dieser signifikante Unterschied in der geschätzten Persistenz weist darauf hin, dass das bivariate Ausgangsmodell die hohe Persistenz im Volatilitätsprozess nicht nachbilden kann.

Das erweiterte Modell (4.45)-(4.48) wird von Liesenfeld (2001) in zwei Spezifikationen jeweils mit und ohne Parameterrestriktion $\tilde{\alpha}_\psi = 0$ geschätzt.

In der ersten Spezifikation wird $\tilde{\kappa}$ gleich Null gesetzt. Dieser Parameter misst die Auswirkung der vergangenen Preisänderungen auf ω_t , bei $\tilde{\kappa} = 0$ ist die Informationssensibilität symmetrisch. Alle geschätzten Parameter erscheinen plausibel und sind meistens signifikant. Vom besonderen Interesse sind die Parameter δ_ω und $\tilde{\nu}_\omega$, die das stochastische Verhalten von ω_t beschreiben. Für beide untersuchten Aktien sind δ_ω und $\tilde{\nu}_\omega$ signifikant größer Null bereits zum Niveau von 1%. Das ist eine Verbesserung gegenüber dem Ausgangsmodell, in welchem $\delta_\omega = \tilde{\nu}_\omega = 0$. Die Analyse der quadrierten Residuen der Preisänderungen $z_{r,t}^2$ zeigt, dass das erweiterte Modell eine bessere Erklärung der Dynamik der Volatilität liefert, als das Ausgangsmodell mit der konstanten Sensibilität. Für beide Aktien ist der Parameter $\tilde{\alpha}_\psi$, der die relative Wichtigkeit des Faktors ω_t für die Sensibilität gegenüber den beiden Arten von Schocks misst, nicht signifikant größer als Null. Um die Auswirkungen des Parameters $\tilde{\alpha}_\psi$ zu untersuchen, schätzt Liesenfeld (2001) die erste Spezifikation mit der Restriktion $\tilde{\alpha}_\psi = 0$. Die Ergebnisse sind fast identisch mit den

Schätzern der ersten Spezifikation ohne Restriktion. Dies bedeutet, dass sich der Faktor ω_t nur auf die gemeinsamen Schocks $\sigma_{\phi,t}^2$ auswirkt. Da das Handelsvolumen nur von $\sigma_{\psi,t}^2$ und nicht von $\sigma_{\phi,t}^2$ abhängt, bedeutet das, dass der Faktor ω_t für das Handelsvolumen eine vernachlässigbare Bedeutung hat, im Gegensatz zur Bedeutung dieses Faktors für die Volatilität der Preisänderungen. Der Vergleich der Schätzer für δ_ω und δ_λ zeigt, dass der Prozess ω_t , der hauptsächlich die Volatilität beeinflusst, eine hohe Persistenz mit δ_ω von 0,996 bzw. 0,987 aufweist. Der Prozess λ_t , der die Ankunft der Informationen beschreibt und sowohl die Volatilität als auch das Handelsvolumen beeinflusst, weist eine geringere Persistenz mit δ_λ von 0,654 bzw. 0,649 auf. Diese Ergebnisse lassen vermuten, dass die *short-run* Komponente der Volatilität mit dem Prozess der Ankunft der Informationen verbunden ist, wogegen die *long-run* Komponente von der Informationssensibilität der Händler abhängt, insbesondere der Sensibilität auf die für alle gleichen Schocks.

In der zweiten geschätzten Spezifikation des erweiterten Modells (4.45)-(4.48) erlaubt Liesenfeld (2001) ein asymmetrisches Verhalten der Informationssensibilität. Das wird erreicht indem der bedingte Erwartungswert von ω_t von den vergangenen Preisänderungen in Abhängigkeit gesetzt wird. Gleich wie in der ersten Spezifikation, ist der Parameter $\tilde{\alpha}_\psi$ statistisch nicht signifikant, das heißt die Veränderungen der Sensibilität gegenüber den spezifischen Schocks und deren Auswirkungen auf die Volatilität und das Handelsvolumen bleiben minimal. Der Parameter $\tilde{\kappa}$, der die Auswirkung der vergangenen Preisänderungen auf den bedingten Erwartungswert $E(\omega_t | \omega_{t-1}, r_{t-1})$ misst, ist für beide Aktien statistisch signifikant bereits zum Niveau von 1%. In der Varianz von r_t ist der Koeffizient β_1 des Schocks $\sigma_{\phi,t}^2$ positiv, was bedeutet, dass die Volatilität bei fallenden Preisen höher ist als bei steigenden beim symmetrischen Verhalten des Handelsvolumens.

Liesenfeld (2001) untersucht mittels Regression, welcher Prozess von ω_t und λ_t in welchem Ausmaß die historischen Werte der absoluten Preisänderungen erklären bzw. die künftigen Werte vorhersagen kann. Der Erklärungsgehalt der Variablen wurde immer höher ermittelt als der Prognosegehalt, was auf die gefilterten Schätzer für den Prognosegehalt zurückzuführen ist. Die Ergebnisse von Liesenfeld (2001) zeigen, dass der Informationsprozess weniger persistent ist und mehr Störungen aufweist als der Sensibilitätsprozess, was die Prognose

erschwert. Der Erklärungsgehalt und der Prognosegehalt des Informationsprozesses wurden wesentlich niedriger geschätzt, als jene mit dem zusätzlich eingeführten Sensibilitätsprozess. Die gleiche Untersuchung des Handelsvolumens zeigt, dass die Einführung des Sensibilitätsprozesses in das erweiterte Modell keinerlei zusätzlichen Erklärungsgehalt bzw. Prognosegehalt bringt, was auf die nicht signifikanten Werte des Parameters $\tilde{\alpha}_v$ zurückzuführen ist. Alleine durch den Informationsprozess werden ca. 97% des historischen Handelsvolumens erklärt und ca. 26% des künftigen Wertes prognostiziert, was wiederum bestätigt, dass der Informationsprozess viele Störungen aufweist.

Liesenfeld (2001) kommt zum Schluss, dass die Handelsvolumenzeitreihen die Information liefern, die meistens für die kurzfristigen Bewegungen der Volatilität der Preisänderungen ausschlaggebend sind. Der Grund dafür ist, dass der Zusammenhang dieser zwei Prozesse zum großen Teil auf der gemeinsamen Abhängigkeit vom Informationsprozess basiert, der eine niedrige Persistenz und viel Störungen aufweist. Aber diese gemeinsame Abhängigkeit zusammen mit der hoch persistenten händlerspezifischen Komponente in der Volatilität impliziert nur geringfügige Verbesserungen in der auf den historischen Preisänderungen und Handelsvolumen Prognose im Vergleich zur Prognose, die nur auf den historischen Preisänderungen basiert. Andererseits liefern die Handelsvolumenzeitreihen wichtige Informationen für die ex-post Analyse der Volatilität der Preisänderungen. Diese Ergebnisse zeigen, dass die Aufnahme eines weiteren Faktors in das Modell das Handelsvolumen keineswegs von der Volatilität trennt. Während der Informationsprozess meistens den Handelsvolumenprozess bestimmt, beeinflusst der Sensitivitätsprozess eher die Volatilität der Preisänderungen. Liesenfelds (2001) Interpretation dieser Beobachtung ist, dass die Handelsvolumenzeitreihen die *short-run* Bewegungen der Volatilität, die durch den Informationsprozess beeinflusst werden, „herausfiltern“, während die *long-run* Komponente der Volatilität mit dem Prozess der Informationssensitivität der Händler verbunden ist.

4.4 Unterschiedliche statische und dynamische Effekte

In diesem Abschnitt wird die von Mestel (2008) durchgeführte Analyse der kurz- und langfristigen Zusammenhänge in den Renditenzeitreihen und Handelsvolumenzeitreihen präsentiert. Seine Untersuchungen stützen die Ergebnisse von Fleming et al. (2006), wonach die ARCH-Effekte in den Renditezeitreihen durch die Dynamik des Handelsvolumens nicht

erklärt werden können, und zeigen, dass die Volumenzitreihen nur Information über die transitorische Komponente des Varianzprozesses enthalten, während sich die Ereignisse mit persistenter Wirkung auf das Handelsvolumen und die Renditevarianz unterschiedlich auswirken.

4.4.1 Kurzfristiger Zusammenhang

Zur Untersuchung des kurzfristigen Zusammenhangs zwischen Renditevolatilitäten und Handelsvolumina werden von Mestel (2008) drei Modelle aufgestellt. Im Benchmark Modell M1 wird der renditegenerierende Prozess als EGARCH(1,1)-Prozess modelliert:

$$R_t = \gamma_0 + \gamma_1 R_{m,t} + \gamma_2 \sigma_t + u_t; \quad u_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim i.i.d.t(v); \quad (4.50)$$

$$\log(\sigma_t^2) = w + \phi \log(\sigma_{t-1}^2) + \theta_1 \left(\frac{|u_{t-1}|}{\sigma_{t-1}} - E \left[\frac{|u_{t-1}|}{\sigma_{t-1}} \right] \right) + \theta_2 \left(\frac{u_{t-1}}{\sigma_{t-1}} \right). \quad (4.51)$$

In der bedingten Mittelwertgleichung (4.50) ist γ_0 eine Konstante und Parameter γ_1 ist der Beta Faktor. Parameter γ_2 bringt den *risk-return tradeoff* zum Ausdruck, und u_t ist der Zufallsfehler des renditegenerierenden Prozesses.

Um die Leptokurtosis in den Renditen zu berücksichtigen, modelliert Mestel (2008) den Zufallsfehler $\{\varepsilon_t\}$ als Sequenz von *i.i.d.* standardisierten Student- t -verteilten Zufallsvariablen mit ν Freiheitsgraden, Erwartungswert 0 und Varianz 1. Das bedeutet, dass der Zufallsfehler u_t in Abhängigkeit von der Informationsmenge $\mathfrak{I}_t = \{\varepsilon_{t-1}^2, \sigma_{t-1}^2\}$ zum Zeitpunkt t bedingt Student- t -verteilt ist mit Erwartungswert 0 und Varianz σ_t . Die logarithmierte Varianz wird in der Gleichung (4.51) als eine multiplikative Funktion der verzögerten standardisierten Renditeinnovationen u_{t-1} modelliert. Das impliziert eine exponentielle Zunahme in der bedingten Varianz im Falle der Zunahme in den Renditeinnovationen. Volatilitäts-Clustering wird durch den Parameter ϕ zum Ausdruck gebracht. Je näher dieser bei 1 liegt, desto persistenter sind Renditeinnovationen in der bedingten Varianz, das heißt desto stärker beeinflusst die Varianz der Vorperiode den aktuellen Wert. Modellparameter θ_1 misst den Einfluss der verzögerten absoluten, über die Standardabweichung normierten Renditeinnovationen auf die bedingte Varianz. Ist dieser Parameter positiv, bewirken größere

Renditeschocks der Vorperiode höhere Schätzung der Renditevarianz der aktuellen Periode. Parameter θ_2 dient der Abbildung von möglichen asymmetrischen Reaktionen der bedingten Varianz auf verzögerte Innovationen (Leverage-Effekt). Ist dieser Parameter negativ, bedeutet das, dass negative Renditeschocks einen stärkeren Anstieg der Varianz bewirken, als positive.

Das Modell M2 unterscheidet sich von dem Benchmark Modell M1 durch Integration des zeitgleichen Volumen als zusätzlichen Regressor in die EGARCH-Gleichung:

$$\log(\sigma_t^2) = w + \phi \log(\sigma_{t-1}^2) + \theta_1 \left(\frac{|u_{t-1}|}{\sigma_{t-1}} - E \left[\frac{|u_{t-1}|}{\sigma_{t-1}} \right] \right) + \theta_2 \left(\frac{u_{t-1}}{\sigma_{t-1}} \right) + \delta_1 V_t. \quad (4.52)$$

Sollte das Handelsvolumen den Informationsfluss tatsächlich abzubilden vermögen, so sollte es laut Mestel (2008) in der Gleichung (4.52) im Vergleich zur Gleichung (4.51) zum Verschwinden von Persistenz, ausgedrückt durch Parameter ϕ , kommen, und der Parameter δ_1 sollte gleichzeitig positive Werte annehmen.

Im Modell M3 wurden die Volumendaten mittels univariater Box/Jenkins ARMA(p,q)-Methodik in eine prognostizierbare und eine nicht prognostizierbare Komponente aufgespalten. Die nicht prognostizierbare Komponente wird als UV_t bezeichnet. Sie ergibt sich als Residuum $\varepsilon_t = UV_t$ des folgenden ARMA-Modells für das geglättete Volumen V_t (vgl. Gleichung 2.2):

$$\Phi(B)V_t = \Theta(B)\varepsilon_t. \quad (4.53)$$

Die Gleichung des Modells M3 lautet demnach:

$$\log(\sigma_t^2) = w + \phi \log(\sigma_{t-1}^2) + \theta_1 \left(\frac{|u_{t-1}|}{\sigma_{t-1}} - E \left[\frac{|u_{t-1}|}{\sigma_{t-1}} \right] \right) + \theta_2 \left(\frac{u_{t-1}}{\sigma_{t-1}} \right) + \delta_2 UV_t. \quad (4.54)$$

Indem im letzteren Modell nur die nicht prognostizierbare Komponente in die Gleichung für die bedingte Varianz einfließt, wird auf die Erklärung des ARCH-Effekts verzichtet, da die so modellierte Informationsrate als ein unkorrelierter Prozess angesehen wird.

Parameter δ_2 misst den Erklärungsgehalt der nicht prognostizierbaren Informationskomponente.

Die Schätzung der Parameter für alle Modelle wird mittels QML-Methode vorgenommen.

Die für das Benchmark Modell M1 geschätzten Parameter zeigen, dass das EGARCH-Modell gut geeignet ist, den renditegenerierenden Prozess abzubilden. Der geschätzte Persistenzparameter ϕ ist sehr groß, was einen hohen Einfluss der Varianz der Vorperiode auf die der aktuellen Periode bedeutet. Somit wird das Volatilitäts-Clustering belegt. Es findet sich auch eine Bestätigung für den Leverage-Effekt. Signifikant positive Werte für den Parameter θ_1 bringen zum Ausdruck, dass je höher der Renditeschock der Vorperiode ist, desto höher wird die Varianz der aktuellen Periode geschätzt. Für die Hälfte der untersuchten Datenreihen wurde der Parameter θ_2 als signifikant negativ geschätzt, das bedeutet, dass negative Renditeschocks einen stärkeren Anstieg in der Varianz als positive bewirken.

Das Modell M1 vermag den statischen Effekt der Heteroskedastizität nur unzureichend abzubilden. Die geschätzte Kurtosis der standardisierten Renditeinnovationen ist oft größer als in den Rohdaten. Dafür erfasst es den ARCH-Effekt in den Renditezeitreihen sehr gut. Das manifestiert sich darin, dass die LM Teststatistik der quadrierten standardisierten Renditeinnovationen zu *lag* 4 in den meisten Fällen nicht signifikant ist, was bedeutet, dass es keine Korrelationen im zweiten Moment gibt.

Die Schätzwerte für das zweite Modell M2 zeigen eine starke Abnahme der Persistenz in der bedingten Renditevarianz, was bedeutet, dass das in der Modellgleichung berücksichtigte Handelsvolumen das Effekt des Volatilitäts-Clustering gut zu erklären vermag. Der Persistenzparameter ϕ bleibt für die Mehrzahl der Fälle hoch signifikant.

In den meisten Fällen haben Renditeschocks der Vorperiode einen starken positiven Einfluss auf die Varianz der aktuellen Periode. Für den Parameter δ_1 werden hoch signifikante Werte geschätzt. Das bedeutet, dass die Aufnahme des Handelsvolumens in die Gleichung des Modells M2 zur deutlichen Abnahme in der Kurtosis führt. Der statische Effekt der Heteroskedastizität lässt sich durch die Berücksichtigung des Handelsvolumens also gut abbilden. Die LM-Statistik der Renditeschocks ist überwiegend hoch signifikant, was

bedeutet, dass das Handelsvolumen den ARCH-Effekt nicht zu erklären vermag. Diese Ergebnisse stehen im Einklang mit der Arbeit von Fleming et al. (2006). Mestel (2008) kommt zum Schluss, dass die Hypothese, wonach die Dynamik des Handelsvolumens den dynamischen Effekt der Heteroskedastizität erklären kann, zu verwerfen ist.

Diese Schlussfolgerungen lassen vermuten, dass im Handelsvolumen Information über die transitorische Komponente des Varianzprozesses enthalten ist, während die Ereignisse mit persistenter Wirkung auf das Handelsvolumen und die Renditevarianz unterschiedlichen Einfluss haben.

Das stützt die neueren Ansätze zur Mischungsverteilungshypothese (vgl. Andersen (1996), Fleming et al. (2006), Liesenfeld (1998)), wonach die Informationsrate aus heterogenen Komponenten besteht, die unterschiedliche Dynamik aufweisen und verschiedene Auswirkungen auf Renditevarianz und Handelsvolumen haben.

Im Modell M3 wird nur der nicht antizipierte Teil der Information durch das Handelsvolumen repräsentiert. Die Ergebnisse der Schätzung zeigen wiederum, dass das Handelsvolumen vor allem die transitorische Komponente im Prozess der bedingten Renditevarianzen abbildet. Der in diesem Modell M3 geschätzte Persistenzparameter ϕ ist größer, als jener im Modell M2, aber immer noch kleiner, als jener im Benchmark Modell M1. In den meisten Fällen ist die LM-Teststatistik zu *lag* 4 signifikant, was auf die Korrelationen im 2. Moment hindeutet. Die simultan mitgeschätzte Anzahl an Freiheitsgraden ν nimmt zu. Das deutet darauf hin, dass sich die Verteilung der Renditeinnovationen in diesem Modell am ehesten einer Normalverteilung annähert. Das wiederum bedeutet, dass die Berücksichtigung der nicht korrelierten Komponente des Handelsvolumens in der Modellgleichung eine bessere Abbildung extremer Renditeschocks zulässt.

Generell lässt sich sagen, dass eine Berücksichtigung des Handelsvolumens in der Modellierung des Prozesses der bedingten Renditevarianzen eine bessere Anpassung an die empirischen Daten ermöglicht. Die vorliegende Analyse von Mestel (2008) bestätigt die Vermutung, dass jener Komponente der Information mit transitorischer Wirkung auf den Varianzprozess, die durch das Handelsvolumen in der Mischungsverteilungshypothese abgebildet wird, eine große Bedeutung zukommt.

4.4.2 Langfristiger Zusammenhang

Nach der Analyse des kurzfristigen Zusammenhangs zwischen Renditevolatilitäten und Handelsvolumen, die im vorangegangenen Abschnitt präsentiert wurde, hat Mestel (2008) eine Untersuchung des bivariaten *long memory* Verhaltens dieser beiden Zeitreihen auf die möglichen Gemeinsamkeiten hin vorgenommen.

Die Handelsvolumenzeitreihen zeigen sehr langfristige Abhängigkeiten und werden daher von Mestel (2008) als ein fraktional integrierter Prozess angesehen (vgl. Abschnitt 2.3). Baillie et al. (1996) stellen einen FIGARCH-Prozess vor, der eine flexible Modellierung der beobachteten Dynamik in den bedingten Renditevarianzen ermöglicht. Der FIGARCH-Prozess kombiniert viele Eigenschaften eines fraktional integrierten Prozesses für den Mittelwert mit den Eigenschaften des GARCH-Prozesses für die bedingte Varianz:

$$\phi(B)(1-B)^d \varepsilon_1^2 = \omega + [1 - \beta(B)]v_t. \quad (4.55)$$

Dabei ist $\{\varepsilon_t\}$ ein diskreter ARCH-Prozess, d ist der fraktionale Differenzparameter und $\{v_t\}$ sind die Innovationen in der bedingten Varianz, die als $v_t = \varepsilon_t^2 - \sigma_t^2$ definiert sind. Im Speziellen, impliziert der FIGARCH-Prozess ein langsames hyperbolisches Abschwingen in den verzögerten quadrierten Innovationen in der bedingten Varianz.

Laut der Annahme der Mischungsverteilungshypothese werden die beiden Prozesse von der latenten Informationsrate gesteuert. Wenn diese als fraktional integrierter Prozess spezifiziert wird, folgt ein identisches *long memory* Verhalten für den Varianzprozess wie den Handelsvolumenprozess. Bollerslev und Jubinski (1999) bestätigen mit ihrem semiparametrischen Ansatz, dass das Handelsvolumen und die Volatilität einen ähnlichen Grad an fraktionaler Integration besitzen. Fleming und Kirby (2006) analysieren Hochfrequenz-Daten und kommen zum Schluss, dass das Handelsvolumen und die Volatilität zwar ein *long memory* Verhalten aufweisen, dieses aber nicht immer identisch ist. Die von Mestel (2008) durchgeführte Untersuchung des kurzfristigen Zusammenhangs führt zur ähnlichen Schlussfolgerung, da der GARCH-cum-volume-Ansatz die Bedeutung des Handelsvolumens nur für die Erklärung der transitorischen Effekte im Volatilitätsprozess bestätigt hat.

Zur Überprüfung möglicher Gemeinsamkeiten in den *long memory* Verhalten der beiden Zeitreihen stellt Mestel (2008) ein Modell auf, in dem er die fraktionalen Differenzparameter für die quadrierten Rendite- und Volumenzeitreihen vergleicht.

Im ersten Schritt wird mittels der modifizierten *rescaled range statistic* überprüft, ob die entsprechenden Datenreihen *long memory* Verhalten aufweisen. Im zweiten Schritt wird der fraktionale Differenzparameter separat für jede Zeitreihe geschätzt. Dann wird diese univariate Schätzung zu einem multivariaten Ansatz erweitert, der die Antwort auf die Frage liefern kann, ob die fraktionalen Differenzparameter der Prozesse signifikant unterschiedlich oder identisch sind.

Nach der Analyse der Datensätze, die eine genügende Anzahl an Realisationen enthalten, kommt Mestel (2008) zum Ergebnis, dass das *long-memory* Verhalten tatsächlich in allen Fällen vorhanden ist. Das heißt, dass Schocks sowohl in den Renditen als auch in den Aktienmengen auch nach einer beträchtlichen Verzögerung nachwirken.

Nun ist die Frage von Interesse, ob das *long-memory* Verhalten des Volatilitätsprozesses und des Volumenprozesses gleich oder unterschiedlich ist. In der Abbildung 4 werden die geschätzten Differenzparameter der quadrierten Renditen jenen des Handelsvolumens gegenübergestellt.

In den meisten Fällen sind Schätzer für den fraktionalen Differenzparameter des Handelsvolumens größer als jene der quadrierten Renditen. Mestel (2008) kommt zum Schluss, dass das *long-memory* Verhalten dieser zwei Prozesse unterschiedlich ist. Diese Schlussfolgerung findet Bestätigung in den statistischen Daten, wonach die Nullhypothese $H_0: d_1 = d_2$ durchschnittlich in 2 von 3 Fällen zu verwerfen ist.

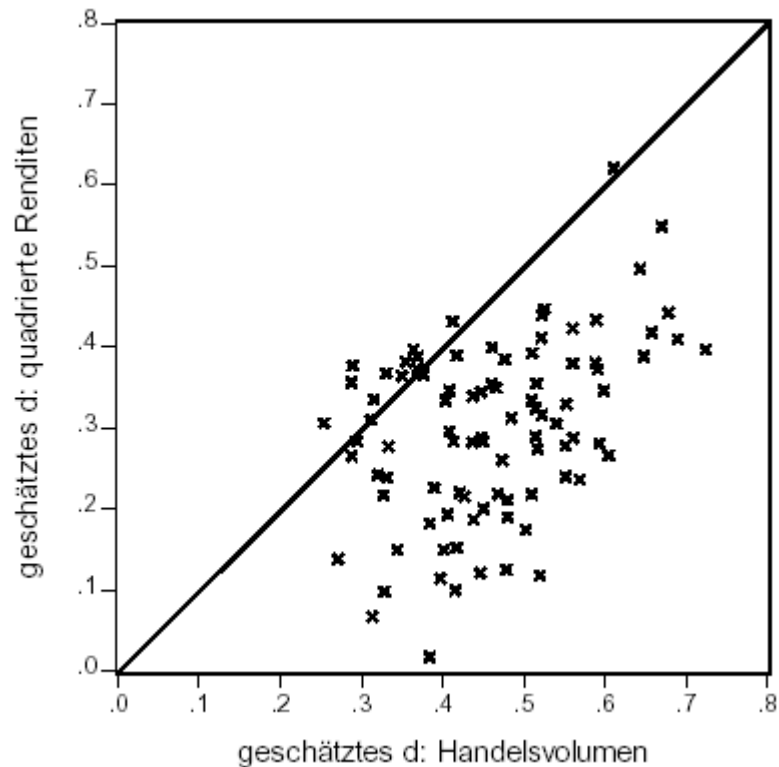


Abbildung 4: Schätzer für den fraktionalen Differenzparameter der quadrierten Renditen und Handelsvolumina ausgewählter Unternehmungen

Quelle: Mestel (2008), Seite 219.

Für die Mischungsverteilungshypothese bedeutet das, dass unterschiedliche Ereignisse Schwankungen im Volatilitätsprozess und im Volumenprozess beeinflussen. Da diese beiden Prozesse annahmegemäß von dem latenten Informationsfluss gesteuert werden, lässt das schlussfolgern, dass der Informationsfluss aus heterogenen Komponenten besteht, die sich auf die beiden Prozesse unterschiedlich stark auswirken.

4.5 Zusammenfassung

Das vierte Kapitel befasste sich mit den möglichen Erklärungen für den Rückgang des Persistenzparameters in bivariaten Modellen. Zuerst wurde die Vermutung von Liesenfeld (1998), Andersen (1996) und Mestel (2008) untersucht, wonach die Informationsrate aus heterogenen Komponenten besteht. Kumar (2008) unterscheidet qualitative Faktoren, die das Handelsvolumen bestimmen und in der Portfoliotheorie und im *behavioural finance* liegen. Fleming et al. (2006) berücksichtigen eine persistente sowie eine nicht persistente

Komponente des Informationsflusses, indem sie mehrere Faktoren in die Spezifikation der Informationsrate aufnehmen.

Danach wurde der Ansatz von Liesenfeld (2001) präsentiert, in dem er neben der latenten Informationsrate die latente Informationssensibilität einführt, welche die Varianz der Preisänderungen mitbeeinflusst. Er kommt zum Schluss, dass nur die kurzfristige Dynamik der Varianz vom Eintreffen der preisrelevanten Informationen bestimmt wird; die langfristige Dynamik ist darauf zurückzuführen, dass die Marktteilnehmer nicht alle eintreffenden Informationen gleich gewichten, das heißt die Sensibilität für die neuen Informationen ist zu verschiedenen Zeitpunkten unterschiedlich.

Zum Schluss wurden in Anlehnung an die empirischen Untersuchungen von Mestel (2008) die statischen und dynamischen Effekte in Renditezeitreihen wie Handelsvolumenzeitreihen erläutert. Während der statische Effekt (*fat tails*) durch die Aufnahme des Handelsvolumens in die Modellgleichung deutlich abnimmt, bleiben die ARCH-Effekte erhalten. Eine mögliche Erklärung dafür ist, dass die Persistenz in beiden Prozessen durch unterschiedliche Faktoren generiert wird. Mestel (2008) unterstützt die Vermutung, dass die Informationsrate aus heterogenen Komponenten besteht, die unterschiedliches dynamisches Verhalten aufweisen. Zur Bestätigung dieser Schlussfolgerung wurde in diesem Kapitel auf die von Mestel (2008) durchgeführte Analyse des *long-memory* Verhaltens der beiden Prozesse eingegangen. Er kommt zum Schluss, dass die Differenzparameter der Handelsvolumina tendenziell größer als jene der quadrierten Renditen sind, was auf den unterschiedlichen Grad an *long memory* hinweist.

5. Zusammenfassung

Die MDH stellt ein empirisch relevantes und flexibles Instrument zur Untersuchung der Entwicklungen auf den Finanzmärkten dar. Mit dieser Hypothese hat Clark (1973) ein Modell erschaffen, das die Leptokurtosis in den Preisänderungszeitreihen erfassen kann. Das wird durch das Mischen der Normalverteilungen mit unterschiedlichen Varianzen erreicht, wobei die latente Informationsrate als Mischungsvariable dient. Neben diesem Standardmodell werden in dieser Arbeit seine Erweiterungen vorgestellt. Zu den dynamischen Erweiterungen zählen das ARV-Modell von Taylor (1986) und die Modelle der ARCH-Klasse. Damit kann eine weitere spezifische Eigenschaft, nämlich das autoregressive Verhalten der Varianz der Preisänderungen, erfasst werden. Die von Liesenfeld (1998) durchgeführte empirische Untersuchung zeigt, dass sowohl die ARV- als auch die ARCH-Modelle die beobachtbare serielle Korrelation in der Varianz der Preisänderungen ähnlich gut erfassen können, wobei das EGARCH- t -Modell die besten Ergebnisse erzielt. Das Problem der Endogenität des Handelsvolumens wird in den bivariaten Erweiterungen berücksichtigt. In diesen Modellen von Tauchen und Pitts (1983) sowie Andersen (1996) wird das Handelsvolumen gemeinsam mit der Volatilität der Preisänderungen von der Informationsrate bestimmt. Die Spezifikation von Tauchen und Pitts (1983) kann als ein statischer Spezialfall der modifizierten MDH von Andersen (1996) gesehen werden. Andersen (1996) und Liesenfeld (1998) führen die Schätzung des modifizierten Modells durch und stellen fest, dass der geschätzte Persistenzparameter signifikant kleiner als in den dynamischen univariaten Modellen ist. Das kann mehrere Gründe haben. Kumar (2008) und Fleming et al. (2006) zeigen, dass die Informationsrate heterogen ist und dass sich Informationsprozesse verschieden auf die Volatilität und Handelsvolumen auswirken. Liesenfeld (2001) führt einen zusätzlichen Faktor in das Modell von Tauchen und Pitts (1983) ein, womit er die zeitabhängige Sensitivität der Händler gegenüber den Informationen nachbilden kann. Er kommt zum Schluss, dass die Volumenzeitreihen mit der *short-run* Dynamik der Volatilität verbunden ist, während die *long-run* Komponente der Volatilität mit der Informationssensitivität der Händler verbunden ist. Mestel (2008) untersucht den kurz- und langfristigen Zusammenhang zwischen dem Volatilitäts- und Volumenprozess und kommt zum Schluss, dass das Handelsvolumen Information über die transitorische Komponente der Volatilität enthält, während der *long-memory* Grad dieser Prozesse unterschiedlich ist.

Literaturverzeichnis

- Andersen, T.G. (1996): Return volatility and trading volume: An informational flow interpretation of stochastic volatility. *The Journal of Finance*, 51, 169-204.
- Baillie, R.T.; Bollerslev, T.; Mikkelsen, H.O. (1996): Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74, 3-30.
- Becker, R. (1998): Die Verallgemeinerte Momentenmethode (Generalized Method of Moments - GMM): Darstellung und Anwendung. *Arbeitspapier Nr. 16*. Online unter: http://www.statoek.vwl.uni-mainz.de/Dateien/Arbeitspapier_Nr_16_GMM.pdf.
- Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307-327.
- Bollerslev, T.; Engle, R. F.; Nelson, D. B. (1994): ARCH models. In: Engle, R. F.; McFadden, D. L. (ed.): *Handbook of Econometrics*, 4, 2959-3038, Elsevier Science B.V.
- Bollerslev, T.; Jubinski, D. (1999): Equity trading volume and volatility: latent information arrivals and common long-range dependencies. *Journal of Business and Economic Statistics*, 17, 9-21.
- Clark, P. K. (1973): A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41, 135-155.
- Danielsson, J.; Richard, J. F. (1993): Accelerated Gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics*, 8, 153-173.
- Engle, R. F. (1982): Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.
- Fama, E.F. (1965): The behaviour of stock-market prices. *Journal of Business*, 38, 34-105.
- Fleming, J.; Kirby, C. (2006): Long memory in volatility and trading volume. *Working Paper*. Online unter: <http://www.ruf.rice.edu/~jfleming/wp/fracint.pdf>.
- Fleming, J.; Kirby, C.; Ostdiek, B. (2006): Stochastic volatility, trading volume, and the daily flow of information. *Journal of Business*, 79, 1551-1590.
- Glosten, L.; Milgrom, P. (1985): Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 13, 71-100.
- Hsieh, D. A. (1989): Modeling Heteroskedasticity in daily foreign-exchange rates. *Journal of Business & Economic Statistics*, 7, 307-317.

- Kumar, A. (2008): Determinants of Stock Trading Volume: Evidence from Indian Stock Markets. Online unter: <http://ssrn.com/abstract=947429>.
- Liesenfeld, R. (1998): *Preise und Handelsvolumina auf Finanzmärkten*. DeutscherUniversitätsVerlag.
- Liesenfeld, R. (2001): A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics*, 104, 141-178.
- Mandelbrot, B.B. (1963): The variation of certain speculative prices. *Journal of Business*, 36, 394-419.
- Mantenga, R. N.; Stanley, H. Eu. (2000): *An introduction to econophysics: correlations and complexity in finance*. Cambridge University Press.
- Mestel, R. (2008): *Handelsvolumen auf Aktienmärkten: Univariate Analysen und kontemporäre Rendite-Mengen-Beziehungen*. Gabler Edition Wissenschaft.
- Nelson, D. B. (1990): ARCH models as diffusion approximations. *Journal of Econometrics*, 45, 7-38.
- Nelson, D. B. (1991): Conditional heteroskedasticity in asset returns : A new approach. *Econometrica*, 59, 347-370.
- Pong, S. E.; Shackleton, M. B.; Taylor, S. J. (2008): Distinguishing Short and Long Memory Volatility Specifications. Online unter: <http://ssrn.com/abstract=879742>.
- Ronning (2003): Korreferat zu: Olaf Hübler: Neuere Entwicklungen in der Mikroökonomie. In: Franz, W.; Ramser, H. J.; Stadler, M. (Hrsg.): *Empirische Wirtschaftsforschung: Methoden und Anwendungen*, 32, 41-49, Mohr Siebeck.
- Schröder, M. (2002): *Finanzmarkt-Ökonometrie: Basistechniken, Fortgeschrittene Verfahren, Prognosemodelle*. Schäffer-Poeschel Verlag.
- Tauchen, G. E.; Pitts, M. (1983): The price variability-volume relationship on speculative markets, *Econometrica*, 51, 485-505.
- Taylor, S. J. (1994): Modeling stochastic volatility: A review and comparative study. *Mathematical Finance*, 4, 183-204.
- Tsay, R. S. (2002): *Analysis of Financial Time Series: Financial Econometrics*. John Wiley and Sons.