

## CDS und Ausfallwahrscheinlichkeitskorrelationen in Krisenzeiten

---

Evidenz für deutsche Kreditinstitute

### **Dissertation**

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften

der Fakultät Statistik  
der Technischen Universität Dortmund  
vorgelegt von

**Carlos Armand Bassi Mbobda**

**Tag der mündlichen Prüfung:** 17. Januar 2013

**Prüfungsvorsitzende:** Prof. Dr. Katja Ickstadt

**Gutachter:** Prof. Dr. Walter Krämer  
Prof. Dr. Rafael Weißbach  
JProf. Dr. Dominik Wied

## Danksagung

Es gibt viele, ohne deren Hilfe diese Dissertation nicht möglich gewesen wäre.

Zuvorderst möchte ich mich bei meinem Doktorvater Rafael Weißbach für seine hervorragende Betreuung und seine nicht endende Unterstützung, die dieses Projekt realisierbar gemacht hat, bedanken. Rafael, danke, dass du mir bereits am Anfang meines Studiums 2008 in Dortmund vertraut hast und trotz der bekannten Schwierigkeiten bei der Online-Betreuung mir die Gelegenheit gegeben hast, an einem aktuellen und interessanten Thema der Finanzökonomie zu arbeiten.

Danke an Walter Krämer für seine wertvolle Anregungen zur Entwicklung dieses Themas im Rahmen des *Postbank Finance Awards 2009* und das Vertrauen in mich und das Projekt, das mir ermöglicht hat, das Projekt an der Fakultät Statistik der TU Dortmund zu schreiben. Danke auch an Dominik Wied für wertvolle Anregungen während der ersten Phase dieser Dissertation.

Ich bin Tim Hanson (Statistics Department, University of South Carolina), Alejandro Jara (Department of Statistics, Pontificia Universidad Católica de Chile), Felix Heinzl (Institut für Statistik, LMU München), Meguy Kuete (Frankfurt School of Finance and Management), Arsene Ntiwa Foudjo (Lehrstuhl für Statistik in den Biowissenschaften, TU Dortmund), Inoncent Agueusop (Global Drug Discovery Statistics, Bayer Pharma AG) für wertvolle Anregungen dankbar.

Mein Dank gilt Christof Lisakowski, Frederike Müller, Natalie Reckmann und Alexander Sommer, studentische Teammitglieder des *Post Bank Finance Awards 2009* der TU Dortmund, für die hervorragende Zusammenarbeit. Ich bedanke mich auch bei Elvis Fiassap Nghonse, Veronika Kretschmer und Sibylle Robens, Arbeitskollegen am IFADO, für Ermutigungen. Ich bin Ann-Josephine Thieme (Lehrstuhl für Statistik, Universität Rostock) für die sprachliche Durchsicht dankbar.

Ich bin meinen Freunden Pierre und Eliane Nguekam, Nadia Noukeu, Eric Kamdem, Flore Nguekam für Ermutigungen dankbar.

Ich bedanke mich schlussendlich bei meiner Familie Pierre Mbobda, Kayo Odette, Jacob und Laure Tame, Janette Yogo, Hugues Didjou Mbobda, Daniel Kenmogne, Ruth Tcheboue Mbobda, Kengmone Bedel, Dadi Messah sowie Giovanni und Christy

Kouemo, für die Liebe und das Vertrauen in mich sowie die Unterstützung meiner Entscheidungen. Ein besonderes Danke schön an Kenmogne Bedel für die finanzielle Unterstützung während dieser Arbeit. Bedel, du hast immer die Verantwortung eines älteren Bruders mir gegenüber übernommen, obwohl du der jüngere bist. Ich habe keinen anderen Weg mich für deine unbedingte Unterstützung zu bedanken.

Carlos Bassi  
Dortmund, 3. Dezember 2012

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung und Zusammenfassung</b>	<b>1</b>
<b>2</b>	<b>Statistische Grundlage der Panel-Regression</b>	<b>7</b>
2.1	ML und REML-Ansatz . . . . .	8
2.1.1	Existenz von ML- und REML-Schätzern . . . . .	9
2.1.2	Eigenschaften von ML- und REML-Schätzern . . . . .	20
2.2	Bayesianischer Ansatz . . . . .	24
2.2.1	Normalverteilung für zufällige Effekte . . . . .	27
2.2.2	Polya-Baum für zufällige Effekte . . . . .	29
<b>3</b>	<b>Korrelation von Ausfallwahrscheinlichkeiten</b>	<b>34</b>
3.1	Finanzökonomische Theorie . . . . .	35
3.2	Korrelation von CDS-Preisen . . . . .	36
3.2.1	Fester Unternehmenseffekt . . . . .	40
3.2.2	Zufälliger Unternehmenseffekt . . . . .	46
3.2.3	Zufälliger Zeiteffekt . . . . .	58
3.2.4	Zufällige Unternehmens- und Zeiteffekte . . . . .	66
<b>4</b>	<b>Material und Auswertung</b>	<b>73</b>
4.1	Material . . . . .	73
4.2	Auswertung . . . . .	83
<b>5</b>	<b>Diskussion</b>	<b>93</b>
	<b>Tabellenverzeichnis</b>	<b>95</b>
	<b>Abbildungsverzeichnis</b>	<b>96</b>
	<b>Abkürzungsverzeichnis</b>	<b>97</b>
	<b>Literatur</b>	<b>100</b>

## 1 Einleitung und Zusammenfassung

Das Risikomanagement im Kreditgeschäft ist im Hinblick auf die Zunahmen von Insolvenzen von immer stärkerer Bedeutung. Die Ertrags- und Risikosteuerung des Kreditportfolios eines Kreditinstituts benötigt deshalb eine starke und stabile Schätzung der Verlustverteilung mit dem Schwerpunkt auf dem *Credit Value at Risk* (CreditVaR). Inhaltlich stellt der CreditVaR eine Abschätzung dar, um welchen Betrag die Verluste aus Kreditrisiken die über die Marge einkalkulierten, erwarteten Standardrisikokosten innerhalb eines Jahres potenziell übersteigen könnten. Mathematisch-statistische Verfahren zur Verlustabschätzung haben in den letzten Jahren stark an Bedeutung gewonnen. Vor allem zur Bewertung von Einzelengagements wurden Ratingsysteme – und die damit verbundene Schätzung der Ausfallwahrscheinlichkeiten (*Probabilities of Default*) – auf statistischen Modellen entwickelt und im Kreditgeschäft eingesetzt. Die Ausfallwahrscheinlichkeit bezieht sich auf die Insolvenz des Referenzschuldners.

In zunehmendem Maße hat sich der Bedarf im Kreditrisikomanagement auf statistische Analysen des Kreditportfolios fokussiert. Eine spannende Frage ist hier: Wie setzt sich die Verlustverteilung aus Krediten in Krisenzeiten zusammen? Dabei stellt die Korrelation von Ausfallwahrscheinlichkeiten eine Herausforderung bei der statistischen Modellierung dar. Diese Schlüsselgröße muss seit dem im Januar 2007 in Kraft getretenen Basel II zur Bestimmung der Kapitalanforderung in Betracht gezogen werden. Als Begründung wurde vom Basler Ausschuss für Bankenaufsicht unter anderem die damals unzureichende Fähigkeit genannt, im Rahmen solcher Modelle die Risikokorrelationen zwischen den Schuldnern zu ermitteln und einzubeziehen. Die korrekte Quantifizierung von Abhängigkeiten zwischen Ausfallwahrscheinlichkeiten beziehungsweise Bonitätsveränderungen stellt eine Kennzahl zur adäquaten Bestimmung der Verlustverteilung eines Kreditportfolios dar (vergleiche Hamerle und Rösch (2003)). Die Korrelation der Ausfallwahrscheinlichkeiten hat somit einen entscheidenden Einfluss auf den CreditVaR, der wiederum das knappe ökonomische Kapital beeinflusst, da die Differenz zwischen dem CreditVaR und dem erwarteten Verlust das ökonomische Kapital deckt (vergleiche Rosenow und Weißbach (2009)).

Es bietet sich an, die Ausfallwahrscheinlichkeitskorrelationen durch die Korrelationen zwischen Aktienrenditen der Unternehmen zu schätzen (vergleiche Forbes und Rigobon (2002) sowie Hamerle und Rösch (2003)). Die daraus entstehenden Korrelationen sind zwar nützlich, erreichen aber schnell ihre Grenzen in der Risikoabschätzung eines Kreditportfolios, da manche Kreditereignisse zu unterschiedlichen Bewegungen im Börsenmarkt sowie Markt für Kreditderivate führen können.

## 1 Einleitung und Zusammenfassung

Während eine Erhöhung des Fremdkapitals beispielsweise zu einem höheren Kreditrisiko führt, was wiederum teure Preise für Kreditderivate mit sich bringt, führt diese Fremdfinanzierung tendenziell eher dazu, dass die Aktien im Wert steigen (vergleiche Jorion und Zhang (2006)). Sinnvolle Instrumente für die Schätzung der Ausfallwahrscheinlichkeitskorrelationen sind deshalb Kreditderivate. Dieses Vorgehen ist gerechtfertigt, da der Markt für Kreditderivate in den letzten Jahren einen enormen Aufschwung erfahren hat. Dieser Aufschwung ist darin begründet, dass Kreditderivate es den Kreditinstituten ermöglichen das Bonitätsrisiko vom originären Kreditinstitut zu trennen und diese Risikokomponente an den Kapitalmarkt zu übertragen. Wichtigste Kreditderivate für die Messung von Ausfallkorrelationen sind CDS (*Credit Default Swaps*) und CDO (*Collateralized Debt Obligation*) (vergleiche Collin-Dufresne (2009)). CDS haben allerdings den Vorteil, dass sie sich kaum durch Illiquidität beeinflussen lassen (vergleiche Longstaff u.a. (2005)). Innerhalb des Marktes für Kreditderivate stellen CDS die wichtigsten, verbreitetsten und einfachsten Produkte dar. Sie verfügen über den größten Marktanteil; im Jahr 2006 und 2008 betrug er rund ein Drittel in diesem Markt (vergleiche Collin-Dufresne (2009)).

CDS können als Versicherungen gegen einen Kreditausfall aufgefasst werden. Dabei überträgt ein Kreditgeber ein Kreditrisiko an einen Sicherungsgeber gegen einen entsprechend bemessenen Preis, der periodisch gezahlt werden kann. Im Fall des Kreditausfalls des Kreditnehmers garantiert der Sicherungsgeber die Ausgleichzahlung. Der zu zahlende CDS-Preis (*spread*) wird in Basispunkten des Buchwerts der Versicherungssumme ausgedrückt (vergleiche Abbildungen 1 und 2) und bezieht sich dabei auf ein Jahr – auch wenn die anteiligen Zahlungen meist halb- oder vierteljährlich erfolgen – und steht in einem engen Zusammenhang mit der Ausfallwahrscheinlichkeit sowie der Bonität des Kreditnehmers für den jeweiligen Einjahreszeitraum (vergleiche Bluhm u.a. (2003)).

Die am Kreditderivate-Markt aktiv agierenden Teilnehmer sind vor allem Kreditinstitute, Versicherungs- und Industrieunternehmen. Dabei sind Banken die bedeutendste Gruppe von Marktakteuren. Im Jahr 2006 stellten sie 59 % der Sicherungsnehmer und 44 % der Sicherungsgeber. Außerdem sind sie Netto-Sicherungsnehmer und daher überwiegend am Verkauf von Kreditrisiken interessiert (vergleiche Wagner, 2007, S. 18f.). In dieser Arbeit werden Daten aus dem Zeitraum der Bankenkrise, also von Juli 2007 bis Januar 2009, betrachtet. Es interessieren vor allem die deutschen Kreditinstitute und unter diesen die 20 größten. Allerdings liegen nur Werte zu 13 dieser Banken vor. Sechs von ihnen – die Deutsche Bank AG, die Commerzbank AG, die Bayerische Hypo- und Vereinsbank AG (HypoVereinsbank), die Kreditanstalt für Wiederaufbau AöR (KfW), die WestLB AG und die Dresdner Bank AG – zählen

## 1 Einleitung und Zusammenfassung

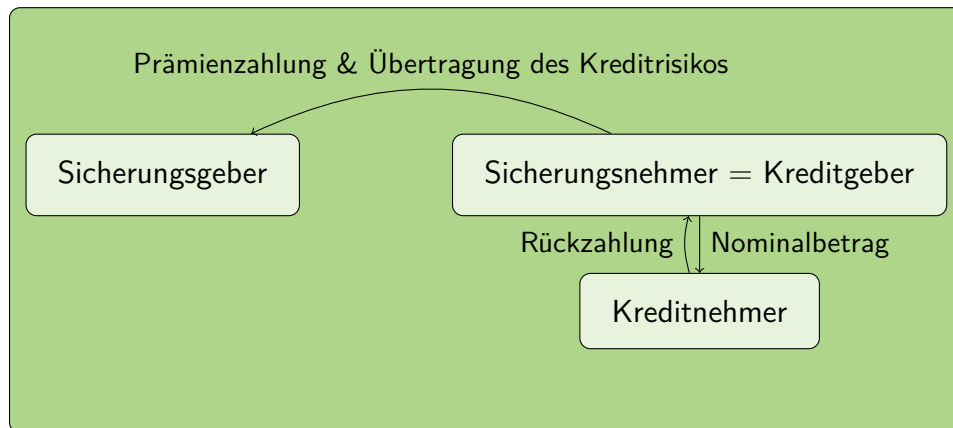


Abbildung 1: Zahlungsströme eines *Credit Default Swaps* ohne Kreditausfall. (Quelle: Bassi Mbobda u.a., 2009, S. 11f.)

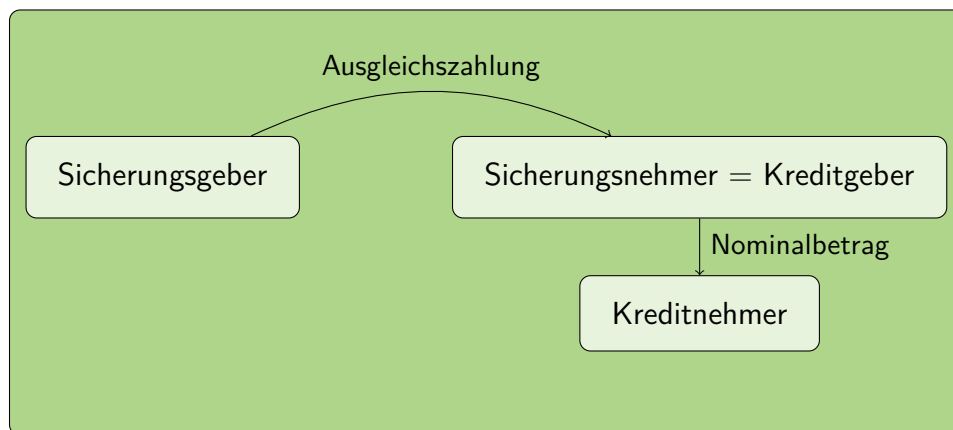


Abbildung 2: Zahlungsströme eines *Credit Default Swaps* bei Kreditausfall. (Quelle: Bassi Mbobda u.a., 2009, S. 11f.)

auch international zu den führenden Akteuren im Kreditderivate-Markt. Die Deutsche Bank betrachtet sich selbst als einen der wichtigsten Marktteilnehmer neben Morgan Stanley, Goldman Sachs, JP Morgan Chase und UBS (vergleiche Wagner, 2007, S. 19).

Großes Gewicht wird der Modellierung von CDS-Preisen der letzten Jahre für das Verständnis sowie das Management von Kreditportfolios beigemessen (vergleiche beispielsweise Jorion und Zhang (2006), Cariboni u.a. (2011), Zhang u.a. (2012) sowie Feldhütter und Nielsen (2012)). Dabei geht es vor allem um die Korrelation von Ausfallwahrscheinlichkeiten, welche wiederum Ausgang des Copula-Modells

## 1 Einleitung und Zusammenfassung

sind (vergleiche beispielsweise Li (2000), McNeil u.a. (2005) und Collin-Dufresne (2009)). Allerdings leidet der Copula-Ansatz unter anderem an den Mängeln, dass dieses Modell an sich statisch sowie dessen Korrelation kein guter Indikator für die Korrelation ist (vergleiche Collin-Dufresne (2009)).

Cherubini u.a. (2011), Creal u.a. (2011) sowie Creal u.a. (2012) haben daher weitere Ansätze wie das verallgemeinerte Autoregressive-Score-Modell oder das dynamische Copula-Modell für die Schätzung von Ausfallwahrscheinlichkeitskorrelationen vorgeschlagen. In dieser Arbeit wird versucht, weitere Modelle zu entwickeln, welche die Abhängigkeitsstruktur von CDS-Preisen besser abbilden können. Parallel zu dem Ansatz von Rosenow und Weißbach (2009) wird die Korrelation von Ausfallwahrscheinlichkeiten mittels Panel-Regressionen bestimmt, wobei nicht Insolvenzraten, sondern CDS-Preise verwendet werden.

In dieser Arbeit wird als neuer Ansatz vorgeschlagen, die zufälligen Effekte durch eine Polya-Baum-Mischung-Verteilung und die Fehler durch einen autoregressiven Prozess zu modellieren. Die Existenz von Maximum-Likelihood- und restringierten Maximum-Likelihood-Schätzern – unter Annahme vorliegender  $AR(1)$ -Fehler – wird wahrscheinlich erstmalig in dieser Arbeit untersucht. Ferner wird eine Formel für die Berechnung der Korrelation im Kontext der Panel-Regressionen, die stilisierten Fakten der Abhängigkeit gerecht wird, hergeleitet. Außerdem werden zwei Ansätze zu einer robusten Parameterschätzung angewendet: der restringierte Maximum-Likelihood-Ansatz sowie die Bayesianische Methode.

Um die Abhängigkeitsstruktur der größten deutschen Kreditinstitute in Krisenzeiten modellieren zu können, werden balancierte Panel-Regressionen von CDS-Preisen durchgeführt. Dabei werden vier Arten von Panel-Regressionen unterschieden: Modelle mit festem Effekt, mit zufälligem Unternehmenseffekt, mit zufälligem Zeiteffekt und mit zufälligen Unternehmens- und Zeiteffekten. Für jedes Modell untersuchen wir zwei Spezifikationen: ein statisches und ein dynamisches Modell. Bei den dynamischen Modellen wurde außerdem untersucht, inwieweit das Marktportfolio (iTraxx) einen Einfluss hat. Die Modelle mit zufälligen Effekten wurden weiter unterteilt in solche mit  $AR(1)$ -Prozess im Fehler und solche ohne.

Für die Analyse der Querschnitt- und Längsschnittkorrelation mithilfe der Panel-Regressionen kann davon ausgegangen werden, dass die Störgrößen mit zufälligen unternehmensspezifischen und/oder zeitspezifischen Effekten modelliert werden können. Die zufälligen unternehmensspezifischen und zeitspezifischen Effekte bewirken, dass die abhängige Variable über die Zeit beziehungsweise über Unternehmen korreliert ist. Die Idee ist, dass wenn einer der zufälligen Effekte tatsächlich keinen Einfluss auf die abhängige Variable hat, deren Varianz auf etwa null ge-



## 1 Einleitung und Zusammenfassung

geschätzt werden sollte. Wir stellen nach Anwendung von 36 Modellen auf die Daten fest, dass

1. etwa 97 % der heutigen CDS-Preise auf die gestrigen zurückzuführen sind.
2. Modelle mit  $AR(1)$ -Fehlern gegenüber Modellen ohne die Anpassung besser – im Sinn des AIC sowie der Längsschnittkorrelation – sind. Diese Modelle führen leider zu einer Inflation – insbesondere bei den statischen Modellen – der Fehlervarianz, was wiederum bewirkt, dass die Querschnittkorrelation Null wird.
3. wenn im Modell die verzögerten CDS-Preise in Betracht gezogen werden, die zufälligen unternehmensspezifischen Effekte auf etwa Null geschätzt werden. Dies deutet darauf hin, dass die verzögerte endogene Variable die Information über mögliche unternehmensspezifische – erklärende – Variable enthält. Die Implikation dieses Befundes auf die Untersuchung ist, dass wenn einige Variablen in der Modellierung unberücksichtigt bleiben, diese wahrscheinlich nicht unternehmensspezifisch sondern zeitspezifisch wirken. Dies steht im Zusammenhang mit Norden und Wagner (2008), die bereits angaben, dass die wesentliche und relevante Informationen für das Kreditgeschäft in den CDS-Preisen enthalten sind.
4. die Ausfallkorrelation über Banken und/oder über die Zeit abhängig davon ist, ob das untersuchte Modell statisch oder dynamisch ist. Für ein statisches Modell – ohne  $AR(1)$ -Fehler – bleiben sowohl die Querschnitt- als auch die Längsschnittkorrelation konstant unabhängig von der Anzahl der Tage zwischen Messungen. Einem solchen Ergebnis widersprechen dynamische Modelle, bei denen die Korrelation mit steigendem Abstand zwischen den Handelstagen sinkt. Dies scheint realistischer zu sein, da die empirische Autokorrelation mit größer werdendem Abstand zwischen Messungen sinkt. Daher sind dynamische Modelle bei der Analyse der Korrelationen statischen vorzuziehen.
5. die Spezifikation eines Modells einen starken Einfluss auf den Verlauf der Korrelation hat.

Collin-Dufresne u.a. (2001) sowie Zhang u.a. (2009) konnten 25 % beziehungsweise 77 % der monatlichen Spreadveränderung im Corporate-Bond beziehungsweise im CDS auf die Einflussgrößen (unternehmensspezifische sowie makroökonomische Variablen) zurückführen. Interessant bei Collin-Dufresne u.a. (2001) ist, dass auch sie über eine Hauptkomponentenanalyse zu dem Ergebnis kamen, dass die Residuen ihrer Regression von einem einzelnen Faktor beeinflusst werden, der nicht

## *1 Einleitung und Zusammenfassung*

unternehmensspezifisch und bislang nicht identifiziert ist. Auch in dieser Arbeit bleibt die Identität des Faktors, durch den die Querschnittkorrelation erklärt werden kann, offen.

Als Modell für die Schätzung von Ausfallwahrscheinlichkeitskorrelationen wird ein dynamisches Modell mit zufälligen zeitspezifischen Effekten (ohne Marktindex als Kovariable) vorgeschlagen. Eine Alternative wäre hier ein dynamisches Modell mit zufälligen zeitspezifischen Effekten und mit dem Marktindex als Kovariable. Diese Modelle besitzen eine perfekte Anpassung an die empirisch beobachteten Längsschnittkorrelationen. Außerdem ergibt sich ausschließlich bei diesen Modellen auch theoretisch ein plausibles Zusammenwirken der Längs- und Querschnittskorrelation.

Die resultierenden Ergebnisse zeigen, dass die CDS-Preise einzelner Banken sowohl untereinander, als auch über die Zeit – und zwar mit einer Querschnittkorrelation von 0,36 sowie einer Ein-Tages-Abhängigkeit von 0,70 – signifikant und hoch korreliert sind. Dies deutet einerseits auf einen höheren Ansteckungseffekt über Banken sowie andererseits auf einen starken linearen Zusammenhang zwischen aufeinanderfolgenden CDS-Preisen innerhalb einer Bank in Krisenzeiten hin. Die relativ höhere Längsschnittkorrelation korrespondiert vermutlich mit der Erfahrung vieler Unternehmen, bei denen bei Schieflagen und Insolvenzen Managementfehler und Unregelmäßigkeiten als Hauptursachen identifiziert werden konnten (vergleiche Hamerle und Rösch (2003)). Einschränkend muss darauf hingewiesen werden, dass die Resultate lediglich auf den Daten von 13 statt aller 20 Banken basieren.

Im weiteren Verlauf der Arbeit werden zunächst in Kapitel 2 die statistischen Grundlagen für das Problem vorgestellt. Im Kapitel 3 wird der Zusammenhang zwischen der Ausfallwahrscheinlichkeitskorrelation und der Korrelation von CDS-Preisen erstellt. Desweiteren werden CDS-Preise mittels der Panel-Regressionen modelliert, Lemmata für die Berechnung der Korrelation formuliert und bewiesen. In Kapitel 4 werden die Daten beschrieben und anschließend die Modelle auf die Daten angewendet. In Kapitel 5 werden die Ergebnisse diskutiert.

## 2 Statistische Grundlage der Panel-Regression

Weil CDS-Preise mehrerer Unternehmen im Zeitablauf analysiert werden sollen, liegt die typische Situation vor, in der die Querschnitt- und Längsschnittkorrelation untersucht werden müssen. Zweckvolle statistische Modelle für diese Analyse bieten die Modelle der Panel-Regression. Hamerle und Rösch (2003) sowie Rose now und Weißbach (2009) untersuchten bereits die Querschnittsabhängigkeit aus Bonitätskennziffern beziehungsweise amt-statistischen Insolvenzzeitreihen mit einem derartigen – aber sehr einfachen – Modell. Es wird hier das folgende Modell betrachtet

$$y_{nt} = \mu + \sum_{k=1}^{m-1} \delta_k x_{ntk} + v_{nt}, \quad n = 1, \dots, N, t = 1, \dots, T. \quad (1)$$

Dabei repräsentieren  $y_{nt}$  die Marktindizes des Unternehmens  $n$  in der Zeitperiode  $t$ .  $\mu$  ist das mittlere Niveau betrachteter Unternehmen auf den Markt. Die  $m - 1$  unabhängige Variablen  $x_{..k}$  repräsentieren Einflußgrößen wie beispielweise regionabhängiger Marktindex oder unternehmensspezifische Variablen.  $\mu, \delta_k, k = 1, \dots, m - 1$  sind in der Literatur als *feste Effekte* eingeführt. Sie geben den Einfluss der unabhängigen Variablen auf die abhängigen Variablen an.  $v_{nt}, n = 1, \dots, N; t = 1, \dots, T$  werden hier als Störgrößen angesehen. Stellt man das Modell (1) in Matrix-Vektor-Schreibweise dar, so lautet es:

$$y = X\delta + v \quad (2)$$

wobei

- $y = (y_{11}, \dots, y_{1T}, \dots, y_{N1}, \dots, y_{N \cdot T})' \in \mathbb{R}^{N \cdot T}$ ,  
 $v = (v_{11}, \dots, v_{1T}; \dots; v_{N1}, \dots, v_{N \cdot T})' \in \mathbb{R}^{N \cdot T}$ ,
- $X = (1, x_{..1}, \dots, x_{..m-1}) \in \mathbb{R}^{N \cdot T \times m}$ ,  $x_{..k} = (x_{11k}, \dots, x_{1Tk}, \dots, x_{N1k}, \dots, x_{NTk})'$
- $\delta = (\mu, \delta_1, \dots, \delta_m)' \in \mathbb{R}^m$

Die folgende Struktur für die Störgröße wird in dieser Arbeit betrachtet

$$v = \sum_{i=1}^p Z_i \zeta_i + \zeta_0. \quad (3)$$

Dabei repräsentieren die  $\zeta_i$  zufällige Effekte mit fester Designmatrix  $Z_i \in \mathbb{R}^{N \cdot T \times k_i}$  ( $\sum_{i=1}^{p-2} k_i < N \cdot T$ ).  $\zeta_0$  ist der Vektor für die Reststörgröße. Es wird angenommen, dass

die zufälligen Effekte  $\zeta_i | \delta_{\zeta_i}, \Sigma_{\zeta_i}, i = 1, \dots, p$  mit Erwartungswert  $\delta_{\zeta_i}$  und Varianz  $\Sigma_{\zeta_i}$  verteilt sind,  $\zeta_0 | \Sigma_{\zeta_0}$  gemäß  $N_{N.T}(0, \sigma_0^2 \Sigma_{\zeta_0})$ . Zwischen den zufälligen Effekten sowie zwischen zufälligen Effekten und Reststörgrößen wird eine stochastische Unabhängigkeit unterstellt. Für die Schätzung der Parameter werden wir eine auf Likelihood sowie eine auf der Bayes'schen Regel basierende Methode anwenden. Dabei werden erstens der Maximum Likelihood (ML)- und der restringierte Maximum Likelihood (REML)-Ansatz und zweitens die Bayesianische Methode vorgestellt.

## 2.1 ML und REML-Ansatz

In diesem Abschnitt werden Parameter durch die Optimierung der Likelihood-Funktion geschätzt. Weitere Annahme über Störgrößen und den Parametervektor sollen gemacht werden, um eine vollständige Likelihood-Funktion zu definieren. Da in der finanzwirtschaftlichen Zeitreihen typischen Strukturen eines stochastischen Prozesses öfter in den Reststörgrößen vorkommen, werden wir den Fall betrachten, wo die Reststörgröße die Struktur eines AR(1)-Prozesses aufweist. Dies lässt sich aber auch später in der Modellierung bestätigen (siehe Bassi Mbobda (2010) für nähere Details zum durchgeführten Likelihood-Quotienten-Test). Es wird angenommen, dass die zufälligen Effekte  $\zeta_i$  gemäß  $N_{k_i}(0, \sigma_i^2 I)$ ,  $i = 1, \dots, p$  verteilt sind sowie  $\zeta_0$  gemäß  $N_{N.T}(0, \sigma_0^2 I_N \otimes V(\phi))$  mit

$$V(\phi) = \frac{1}{1 - \phi^2} \begin{pmatrix} 1 & \phi & \phi^2 & \dots & \phi^{T-1} \\ \phi & 1 & \phi & \dots & \phi^{T-2} \\ \phi^2 & \phi & 1 & \dots & \phi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi^{T-1} & \phi^{T-2} & \phi^{T-3} & \dots & 1 \end{pmatrix}, \quad (4)$$

wobei  $\otimes$  das Kronecker-Symbol repräsentiert. Die Positivität der unbekanten Varianzparameter ( $\sigma_0^2 > 0, \sigma_i^2 \geq 0, i = 1, \dots, p$ ) wird in dieser Arbeit vorausgesetzt. Die natürliche Wahl von  $V(\phi)$  ist die Einheitsmatrix. Wir wollen uns aber nicht auf diesen Fall beschränken, sondern uns alle Fälle anschauen, in denen gilt  $|\phi| < 1$ . Es ergibt sich der nichtkompakte Parameterraum  $\Gamma := \{\vartheta = (\delta, \theta) \in \mathbb{R}^m \times \Theta\}$  mit  $\Theta := \{\theta = (\sigma_0^2, \sigma_1^2, \dots, \sigma_p^2, \phi) \in (0, +\infty) \times [0, +\infty)^p \times (-1, 1)\} \subset \mathbb{R}^{p+2}$ . Desweiteren gilt  $y \sim N_{N.T}(X\delta, \Sigma)$ . Dabei repräsentiert  $\Sigma$  die Varianz-Kovarianz-Matrix und ist wie folgt gegeben

$$Var(y) = \sum_{i=1}^p \sigma_i^2 Z_i Z_i' + \sigma_0^2 I_N \otimes V(\phi) =: \Sigma.$$

## 2 Statistische Grundlage der Panel-Regression

Die Likelihood-Funktion für das Modell (2) sieht folgendermaßen aus

$$L(\vartheta) := L(\vartheta|y) = (2\pi)^{\frac{-N \cdot T}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y - X\delta)' \Sigma^{-1}(y - X\delta)\right\}. \quad (5)$$

Obwohl die Likelihood-Funktion (5) stetig ist, muss ein Maximum nicht existieren, da der Parameterraum  $\Gamma$  nicht kompakt ist. Wir werden uns in den folgenden mit dem Existenzbeweis der ML- und REML-Schätzer im Modell (2) beschäftigen.

### 2.1.1 Existenz von ML- und REML-Schätzern

Das Buch von Rao und Kleffe (1988), die Artikel von Demidenko und Massam (1999), Birkes und Wulff (2003) sowie De Oliveira und Ferreira (2011) sind wahrscheinlich die einzigen Arbeiten mit dem Ziel, Auskünfte über die Existenz von Maximum- und Restringierten Maximum Likelihood Schätzern zu geben. Allerdings gilt die dort gegebene notwendige und hinreichende Bedingung nicht für das vorliegende Problem. Die in dieser Arbeit gegebene Varianzstruktur unterscheidet sich nämlich in vier Punkten von deren:

- (1) Im Buch von Rao und Kleffe (1988) (Kapitel 9) hat die Varianz die folgende Struktur:  $\Sigma = \theta_1 V_1 + \dots + \theta_{p+1} V_{p+1}$ , wobei  $V_i$ ,  $i = 1, \dots, p + 1$  linear unabhängig sind und der Parameterraum  $\Theta = \{\theta : \Sigma \text{ positiv definit}\}$ . Somit dürfen  $\theta_i$ ,  $i = 1, \dots, p + 1$  auch negativ sein. In statistischem Sinn wären solche negativen Varianzparameter unerwünscht.
- (2) Demidenko und Massam (1999) haben sich mit dem Fall  $\phi = 0$  beschäftigt.
- (3) Birkes und Wulff (2003) beschäftigten sich mit gleicher Varianzstruktur und dem gleichen Parameterraum wie Rao und Kleffe (1988)(Kapitel 9). Allerdings setzen sie zusätzlich voraus, dass  $V_i$ ,  $i = 1 \dots p + 1$  nicht Null und nicht negativ-definit sein dürfen,  $V_{p+1} = I$  sowie kommutativ sind; das heißt  $V_{i_1} V_{i_2} = V_{i_2} V_{i_1}$ ,  $1 \leq i_1, i_2 \leq p + 1$ . Schreiben wir  $V(\phi)$  wie folgt um

$$V(\phi) = \frac{1}{1 - \phi^2} \left( I_T + \sum_{t=2}^T \phi^{t-1} H_t \right),$$

wobei  $H_t$  die nebendiagonale Matrix mit Einsen auf der t-ten unteren und

## 2 Statistische Grundlage der Panel-Regression

oberen Nebendiagonale ist,

$$H_t = \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 1 & \ddots & \ddots & \ddots & \ddots & \ddots & 1 \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}.$$

erhalten wir die folgende Struktur für die Varianz

$$\Sigma = \sum_{i=1}^p \sigma_i^2 Z_i Z_i' + \sum_{t=2}^T \sigma_0^2 \frac{\phi^{t-1}}{1-\phi^2} \mathbf{I}_N \otimes H_t + \frac{\sigma_0^2}{1-\phi^2} \mathbf{I}_N \otimes \mathbf{I}_T. \quad (6)$$

Im vorliegenden Fall sind  $H_t, t = 2 \dots T$  zusätzlich nicht definit und die Kommutativität spielt absolut keine Rolle in dieser Arbeit.

- (4) De Oliveira und Ferreira (2011) betrachteten  $\Sigma = \theta V(\psi)$  mit  $V(\psi)^{-1} = \mathbf{I}_{N \cdot T} + \psi H$  und  $\Theta = \{(\theta, \psi) \in (0, +\infty) \times [0, +\infty)\}$ , wobei  $H$  bekannt ist und folgendermaßen aussieht

$$H_{i,j} = \begin{cases} \sum_{i=1}^j g_{ij} & , \text{ für } i = j \\ -g_{ij} & , \text{ für } i \in N_j \\ 0 & , \text{ sonst.} \end{cases}$$

Dabei ist  $g_{ij} = g_{ji} > 0$ .  $N_j$  bezieht sich auf die Menge der Nachbarn von  $j$  ( $j \notin N_j$ ). Sie setzen außerdem voraus, dass gilt  $H \mathbf{1}_{N \cdot T} = 0$ .

Das Modell (3) unterscheidet sich von diesem Modell in mindestens zwei Punkten:

- $p \neq 0$

## 2 Statistische Grundlage der Panel-Regression

- $V(\phi)^{-1} = I_T + H^*$  mit

$$H^* = \begin{pmatrix} 0 & -\phi & 0 & \cdots & \cdots & \cdots & 0 \\ -\phi & \phi^2 & -\phi & \ddots & & & \vdots \\ 0 & -\phi & \phi^2 & -\phi & \ddots & & \vdots \\ \vdots & \ddots & -\phi & \phi^2 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & -\phi & 0 \\ \vdots & & & \ddots & -\phi & \phi^2 & -\phi \\ 0 & \cdots & \cdots & \cdots & 0 & -\phi & 0 \end{pmatrix}.$$

Hier ist deutlich zu sehen, dass gilt  $H^*1_T \neq 0$ .

Das Varianzkomponenten-Modell (3) repräsentiert also eine leichte Erweiterung der soeben genannten Arbeiten. Eine Folge davon ist, dass Ihre Ergebnisse nicht hier anwendbar sind. Im folgenden werden wir versuchen, das vorliegende Modell zu transformieren und die Existenzbedingung diskutieren. Die Prais-Winsten-Transformation werden wir benutzen und den Satz 3.1 aus Demidenko und Massam (1999) auf das transformierte Problem anwenden. Wenden wir nun die Prais-Winsten-Transformation nämlich

$$V(\phi)^{-1} = C'C \text{ mit}$$

$$C = \begin{pmatrix} \sqrt{1-\phi^2} & 0 & \cdots & \cdots & 0 \\ -\phi & 1 & \ddots & & \vdots \\ 0 & -\phi & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\phi & 1 \end{pmatrix}$$

an, so ergibt sich das folgende Modell

$$y^* = X^* \delta + \sum_{i=1}^{p-2} Z_i^* \xi_i + \xi_0^*. \quad (7)$$

Wobei gilt  $y^* = (I_N \otimes C)y$ ,  $X^* = (I_N \otimes C)X$ ,  $Z_i^* = (I_N \otimes C)Z_i$  sowie  $\xi_0^* = (I_N \otimes C)\xi_0$ .

Somit hat  $y^*$  die folgende Varianz-Kovarianz-Matrix

$$\begin{aligned} \Sigma^* &= \sum_{i=1}^{p-2} \sigma_i^2 (I_N \otimes C) Z_i Z_i' (I_N \otimes C') + \sigma_0^2 I_{N \cdot T} \\ &= (I_N \otimes C) \Sigma (I_N \otimes C'). \end{aligned} \quad (8)$$

## 2 Statistische Grundlage der Panel-Regression

Da  $y^*$  gemäß  $N(X^*\delta, \Sigma^*)$  verteilt ist, lässt sich die Likelihood-Funktion für das Modell (7) wie folgt ausdrücken:

$$L^*(\vartheta) := L^*(\vartheta|y) = (2\pi)^{\frac{-N \cdot T}{2}} |\Sigma^*|^{\frac{-1}{2}} \exp\left\{-\frac{1}{2}(y^* - X^*\delta)' \Sigma^{*-1} (y^* - X^*\delta)\right\}. \quad (9)$$

Somit besteht einen Zusammenhang zwischen den beiden Likelihood-Funktionen, der im folgenden Lemma zusammengefasst wird.

**Lemma 2.1.**

$$L^*(\vartheta) = (1 - \phi^2)^{-\frac{N}{2}} L(\vartheta) \quad (10)$$

*Beweis.*

$$\begin{aligned} (y^* - X^*\delta)' \Sigma^{*-1} (y^* - X^*\delta) &= (y - X\delta)' \mathbf{I}_N \otimes C' (\mathbf{I}_N \otimes C)^{-1} \Sigma^{-1} \times \\ &\quad (\mathbf{I}_N \otimes C)^{-1} \mathbf{I}_N \otimes C (y - X\delta) \\ &= (y - X\delta)' \Sigma^{-1} (y - X\delta) \end{aligned}$$

und

$$\begin{aligned} |\Sigma^*| &= |(\mathbf{I}_N \otimes C) \Sigma (\mathbf{I}_N \otimes C')| \\ &= |(\mathbf{I}_N \otimes C')| |(\mathbf{I}_N \otimes C)| |\Sigma| \\ &= |(\mathbf{I}_N \otimes C' C)| |\Sigma| \\ &= |V(\phi)^{-1}|^N |\Sigma| \\ &= (1 - \phi^2)^N |\Sigma|. \end{aligned}$$

□

**Definition 2.2.** Die ML-Schätzer  $\hat{\vartheta}_{ML}$  sowie  $\hat{\vartheta}_{ML}^*$  für  $\vartheta$  im Modell (2) und (3) beziehungsweise (7) sind durch

$$\arg \max_{\vartheta \in \Gamma} L(\vartheta) \text{ beziehungsweise } \arg \max_{\vartheta \in \Gamma} L^*(\vartheta) \text{ zu bestimmen.}$$

**Bemerkung 2.3.** Die Lösung eines Optimierungsproblems ist invariant durch eine streng isotone Transformation der zu optimierenden Funktion. Es folgt daraus, dass gilt

$$\arg \max_{\vartheta \in \Gamma} L(\vartheta) = \arg \max_{\vartheta \in \Gamma} \log L(\vartheta).$$



## 2 Statistische Grundlage der Panel-Regression

Logarithmiert man die ML-Funktionen (5) und (9), ergibt sich

$$l(\vartheta) = -\frac{1}{2} \left[ N \cdot T \ln(2\pi) + \ln |\Sigma| + (y - X\delta)' \Sigma^{-1} (y - X\delta) \right] \quad (11)$$

beziehungsweise

$$l^*(\vartheta) = -\frac{N}{2} \ln(1 - \phi^2) + l(\delta, \theta). \quad (12)$$

Wir werden uns im folgenden Satz mit der hinreichenden und notwendigen Bedingung für die Existenz vom ML-Schätzer im Modell (7) befassen. Davor führen wir die folgenden Notationen ein. Seien  $M = \langle X \rangle$  ( $M^* = \langle X^* \rangle$ ) und  $H = \langle Z \rangle$  ( $H^* = \langle Z^* \rangle$ ) die durch  $X$  (beziehungsweise  $X^*$ ) und  $Z = [Z_1, \dots, Z_p]$  (beziehungsweise  $Z^* = [Z_1^*, \dots, Z_p^*]$ ) aufgespannten Vektorräume, sodass gilt  $M + H = \langle X, Z \rangle$  beziehungsweise  $M^* + H^* = \langle X^*, Z^* \rangle$ .  $P_{(M^*+H^*)^\perp}$  bezeichnet die orthogonale Projektion von  $\mathbb{R}^{N \cdot T}$  auf  $(M^* + H^*)^\perp$ . Wir definieren außerdem  $S_{X^*, Z^*} := y^{*\prime} P_{(M^*+H^*)^\perp} y^*$ .

**Satz 2.4.** Sei  $y^*$  wie im Modell (7). Der ML-Schätzer  $\vartheta_{ML}^* = (\hat{\delta}_{ML}^*, \hat{\theta}_{ML}^*)$  existiert genau dann, wenn gilt  $y^* \notin \langle X^*, Z^* \rangle$ .

Genau gesagt:

1. Gilt  $y^* \notin \langle X^*, Z^* \rangle$ , dann folgt  $l^*(\vartheta) \geq N \cdot T \ln(S_{X^*, Z^*}) - N \cdot T \ln N \cdot T - N \cdot T$ , und  $l^*(\vartheta)$  erreicht sein Maximum in  $\Gamma$
2. Gilt  $y^* \in \langle X^*, Z^* \rangle$ , dann ist  $+\infty$  das Supremum von  $l^*(\vartheta)$  in  $\Gamma$

**Bemerkung 2.5.** Für  $y^* \notin M^* + H^*$  ist  $\ln(S_{X^*, Z^*})$  wohl definiert, da gilt

$$\begin{aligned} y^* \notin M^* + H^* &\iff \forall e \in (M^* + H^*)^\perp, e \bullet y \neq 0 \\ &\iff y^{*\prime} P_{(M^*+H^*)^\perp} y^* > 0. \end{aligned}$$

Dabei repräsentiert  $\bullet$  das Skalarprodukt.

**Beweis.** Der Satz 2.4 lässt sich mit Demidenko und Massam (1999) (Satz 3.1) beweisen. □

Hätten wir  $\hat{\vartheta}_{ML} = \hat{\vartheta}_{ML}^*$ , dann gäbe der Satz 2.4 auch die hinreichende und notwendige Bedingung für die Existenz vom ML-Schätzer in Modell (2) und (3). Leider geht es in diesem Fall nicht, wie das folgende Lemma zeigt.

**Lemma 2.6.** Existieren  $\hat{\vartheta}_{ML}$  sowie  $\hat{\vartheta}_{ML}^*$ , dann gilt

$$\hat{\vartheta}_{ML} = \hat{\vartheta}_{ML}^* \iff \phi = 0.$$

## 2 Statistische Grundlage der Panel-Regression

**Beweis.** Annahme:  $\hat{\vartheta}_{ML} = \hat{\vartheta}_{ML}^* = \chi$  und  $\phi \neq 0$

$$0 = \frac{\partial}{\partial \phi} L^*(\chi) = N \hat{\phi}_{ML} (1 - \hat{\phi}_{ML}^2)^{-\frac{N+2}{2}} L(\chi) + (1 - \hat{\phi}_{ML}^2)^{-\frac{N}{2}} \frac{\partial}{\partial \phi} L(\chi).$$

Da gilt  $\frac{\partial}{\partial \phi} L(\chi) = 0$  und  $\hat{\phi}_{ML} (1 - \hat{\phi}_{ML}^2)^{-\frac{N+2}{2}} \neq 0$ , folgt  $L(\chi) = 0$ . Widerspruch, da die Likelihood-Funktion (5) weder konstant noch negativ ist.  $\square$

**Lemma 2.7.** Existiert  $\hat{\vartheta}_{ML}^*$ , dann existiert auch  $\hat{\vartheta}_{ML}$ .

**Beweis.** Sei  $\hat{\vartheta}_{ML}^* \in \Gamma$  der ML-Schätzer im Modell (7), dann folgt

$$\begin{aligned} & \forall \vartheta \in \Gamma, L^*(\vartheta) \leq L^*(\hat{\vartheta}_{ML}^*) \\ \iff & \forall \vartheta \in \Gamma, (1 - \phi^2)^{-\frac{N}{2}} L(\vartheta) \leq L^*(\hat{\vartheta}_{ML}^*) \\ \implies & \forall \vartheta \in \Gamma, L(\vartheta) \leq L^*(\hat{\vartheta}_{ML}^*) \\ \implies & \hat{\vartheta}_{ML} \text{ existiert.} \end{aligned}$$

$\square$

**Lemma 2.8.**

$$y^* \notin \langle X^*, Z^* \rangle \iff y \notin \langle X, Z \rangle$$

**Beweis.** Da  $V(\phi)$  positiv definit ist, ist  $I_N \otimes C$  ein Isomorphismus  $\forall |\phi| < 1$ . Es folgt daraus, dass gilt

$$\begin{aligned} y^* \in \langle X^*, Z^* \rangle & \iff \exists (x^*, h^*) \in M^* \times H^* : y^* = m^* + h^* \\ & \iff \exists (m, h) \in M \times H : (I_N \otimes C)y = y^* = (I_N \otimes C)(m + h) \\ & \iff \exists (m, h) \in M \times H : y = m + h \\ & \iff y \in \langle X, Z \rangle. \end{aligned}$$

$\square$

**Korollar 2.9.** Der ML-Schätzer in den Modellen (2) und (3) existiert, wenn gilt

$$y \notin \langle X, Z \rangle.$$

**Bemerkung 2.10.** Demidenko und Massam (1999) konnten die Existenzbedingung von ML-Schätzern in den Modellen (2) und (3) für den Spezialfall  $\phi = 0$  geben. Für deren Fall existiert der ML-Schätzer genau dann, wenn gilt  $y \notin \langle X, Z \rangle$ . Da unsere Untersuchung zeigt, dass dieser Schätzer bei unbekanntem  $\phi$  unter derselben Bedingung existiert, folgt, dass die Existenz der ML-Schätzer bei bekannter Varianz-Kovarianz-Matrix  $\Sigma_{\xi_0}$  die Existenz von ML-Schätzern bei vorliegenden AR(1)-Fehlern impliziert.

## 2 Statistische Grundlage der Panel-Regression

**Lemma 2.11.** Den ML-Schätzer für die Varianz  $\Sigma$  erhält man, indem man die folgende Gleichung löst:

$$\text{spur} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \right) = -y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y, \\ i = 1, \dots, p + 2,$$

wobei gilt

$$\Pi(\theta) = \Sigma^{-1} (I - H(\theta)), \text{ mit } H(\theta) = X \left[ X' \Sigma^{-1} X \right]^{-1} X' \Sigma^{-1}. \quad (13)$$

*Beweis.* Um die logarithmierte ML-Funktion (11) zu maximieren, werden wir sowohl die Rechenregel für die Ableitung der Inverse einer Matrix;

$$\frac{\partial}{\partial \theta_i} \Sigma^{-1} = -\Sigma^{-1} \left( \frac{\partial}{\partial \theta_i} \Sigma \right) \Sigma^{-1} \quad (14)$$

als auch für die Ableitung der logarithmierten Determinante einer Matrix;

$$\frac{\partial}{\partial \theta_i} \ln |\Sigma| = \text{spur} \left( \Sigma^{-1} \frac{\partial}{\partial \theta_i} \Sigma \right) \quad (15)$$

aus Anderson (1970, S. 5f.) nutzen. Leitet man die logarithmierte ML-Funktion (11) nach  $\delta$  und  $\Sigma$  ab und setzt sie gleich Null, ergibt sich;

$$\begin{cases} [X' \Sigma^{-1} X] \delta - X' \Sigma^{-1} y = 0 \\ \text{spur} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \right) - (y - X\delta)' \left[ \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} \right] (y - X\delta) = 0 \end{cases} \quad (16)$$

Löst man die erste Gleichung im System (16), erhält man;

$$\hat{\delta} = \left[ X' \Sigma^{-1} X \right]^{-1} X' \Sigma^{-1} y \quad (17)$$

Setzt man  $\hat{\delta}$  in der zweiten Gleichung, bekommt man;

$$\begin{aligned} \text{spur} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \right) &= (y - H(\theta)y)' \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (y - H(\theta)y) \\ &= y' \Pi(\theta)' \frac{\partial \Sigma}{\partial \theta_i} H(\theta) y \\ &= y' \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Pi(\theta) y - y' H(\theta) \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} y + \\ &\quad + y' H(\theta) \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} H(\theta) \Sigma^{-1} \\ &= -y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y. \end{aligned}$$

□

## 2 Statistische Grundlage der Panel-Regression

Die Gleichungen aus Lemma 2.11 lassen sich im Allgemeinen nicht direkt, sondern mit iterativen Verfahren lösen. Modellansätze zur Lösung sind die auf Expectation Maximization (EM), Newton-Raphson, Simplex, Quasi-Newton und Fisher-Scoring basierenden Algorithmen. In Laird und Ware (1982) wird zur Lösung des Maximierungsproblems der EM-Algorithmus vorgeschlagen. In Lindstrom und Bates (1988) werden jedoch Gründe – wie beispielsweise eine wesentlich geringere Iterationszahl – aufgeführt, die vorrangig den Newton-Raphson-Algorithmus behandeln. Dieser Algorithmus aktualisiert die Parameter durch den folgenden Schritt (vergleiche Song u.a. (2005))

$$\vartheta^l = \vartheta^{l-1} - \left\{ \frac{1}{N} l''(\vartheta^{l-1}) \right\}^{-1} \left[ \frac{1}{N} l'(\vartheta^{l-1}) \right], \quad l = 1, 2, \dots,$$

wobei  $l''(\vartheta^{l-1})$  die Hesse-Matrix der Ableitung zweiter Ordnung von  $l(\cdot)$  in  $\vartheta^{l-1}$  ist.  $\vartheta^0$  repräsentiert dabei den Startwert. Für komplizierte Likelihood-Funktionen schlagen Song u.a. (2005) den Maximazation-by-Parts-Algorithmus (MBP) vor, welcher auf die Aufteilung der Likelihood in zwei Funktionen – wobei eine davon einfacher als die ursprüngliche abzuleiten ist – basiert.

**Lemma 2.12.**

$$E \left[ y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y \right] = - \text{spur} \left( \Pi(\theta) \frac{\partial \Sigma}{\partial \theta_i} \right)$$

*Beweis.*

$$\begin{aligned} y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y &= - y' \Pi(\theta)' \frac{\partial \Sigma}{\partial \theta_i} \Pi(\theta) y \\ &= - \text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Pi(\theta) y y' \Pi(\theta)' \right) \\ &= - \text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (I - H(\theta)) (I - H(\theta)) \Sigma^{-1} y y' \right) \\ &= - \text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (I - H(\theta)) \Sigma^{-1} y y' \right) \end{aligned}$$

## 2 Statistische Grundlage der Panel-Regression

$$\begin{aligned}
 E\left(y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y\right) &= -\text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (I - H(\theta)) \Sigma^{-1} E(yy') \right) \\
 &= -\text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (I - H(\theta)) \Sigma^{-1} (\Sigma + X\delta\delta'X') \right) \\
 &= -\text{spur} \left( \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} (I - H(\theta)) (\Sigma + X\delta\delta'X') \Sigma^{-1} \right) \\
 &= -\text{spur} \left( \Pi(\theta) \left( I + X\delta\delta'X'\Sigma^{-1} \right) \frac{\partial \Sigma}{\partial \theta_i} \right) \\
 &= -\text{spur} \left( \Pi(\theta) \frac{\partial \Sigma}{\partial \theta_i} \right).
 \end{aligned}$$

□

Der ML-Schätzer ( $\hat{\theta}_{ML}$ ) ist aufgrund der Nichtberücksichtigung des Verlustes an Freiheitsgraden oft nicht vorrangig behandelt. Mardia und Marshall (1984), Swallow und Monahan (1984) sowie Zimmerman und Monahan (1989) zeigten anhand von Simulationen, dass ML-Schätzer schwer verzerrt sein können, wenn  $N \cdot T$  nicht genügend groß ist. Abhängig vom Kontext kann die Verzerrung zu gravierenden Folgen führen (vergleiche Cressie und Soumendra, 1993, S. 17f.). Eine Methode, welche diese Unangemessenheit behebt, liefert die REML-Schätzung, eingeführt von Patterson und Thompson (1971) und diskutiert von Harville (1974). Hierbei wird die ML-Methode auf eine lineare Funktion von  $y$  angewandt, welche keinen der festen Effekte beinhaltet. Eine solche lineare Kombination nennt man Fehlerkontrast.

**Definition 2.13.** Sei  $\gamma \in \mathbb{R}^{N \cdot T}$ , sodass gilt  $E(\gamma'y) = 0$  ( $\gamma'X = 0$ ). Dann ist  $\gamma'y$  ein Fehlerkontrast.

**Lemma 2.14.** Sei  $A$  eine  $N \cdot T \times (N \cdot T - m)$  Matrix, sodass gilt

$$I - X(X'X)^{-1}X' = AA' \text{ sowie } I_{N \cdot T - m} = A'A.$$

Dann besitzt  $A$  genau  $(N \cdot T - m)$  linear unabhängige Fehlerkontraste.

**Beweis.** Da der Rang von  $A$   $N \cdot T - m$  beträgt ( $N \cdot T - m = \text{rang}(AA') \leq \text{rang}(A) \leq N \cdot T - m$ ), besitzt  $A$  genau  $N \cdot T - m$  linear unabhängige Vektoren. Außerdem gilt

$$A'X = A'AA'X = A'(I - X(X'X)^{-1}X')X = 0.$$

□

**Lemma 2.15.** Existiert eine Matrix  $A$  mit den Eigenschaften aus Lemma 2.14, dann folgt, dass  $A'y$  und  $\hat{\delta}$  stochastisch unabhängig sind.

## 2 Statistische Grundlage der Panel-Regression

**Beweis.** Da unter Annahme fester Kovarianzparameter  $A'y \sim N_{N \cdot T - m}(0, A'\Sigma A)$  und  $\hat{\delta} = [X'\Sigma^{-1}X]^{-1} X'\Sigma^{-1}y \sim N_m(\delta, [X'\Sigma^{-1}X]^{-1})$  gelten, reicht es zu zeigen, dass  $A'y$  und  $\hat{\delta}$  unkorreliert sind.

$$\begin{aligned}
 \text{kov}(A'y, \hat{\delta}) &= \text{kov}(A'(X\delta + v), \hat{\delta}) \\
 &= E(A'v\hat{\delta}') \\
 &= E\left(A'v([X'\Sigma^{-1}X]^{-1}X'\Sigma^{-1}y)'\right) \\
 &= E\left(A'vy'\Sigma^{-1}X[X'\Sigma^{-1}X]^{-1}\right) \\
 &= E\left(A'v(\delta'X' + v')\Sigma^{-1}X[X'\Sigma^{-1}X]^{-1}\right) \\
 &= A'E(vv')\Sigma^{-1}X[X'\Sigma^{-1}X] \\
 &= A'X([X'\Sigma^{-1}X]) \\
 &= 0.
 \end{aligned}$$

□

Es wird angenommen, dass  $X$  vollen Spaltenrang hat. Die Existenz einer solchen Matrix  $A$  mit den Eigenschaften aus Lemma 2.14 ist dann durch die Spektralzerlegung von  $I - X(X'X)^{-1}X'$  gesichert. Nun wird ein Vektor  $u$  wie folgt definiert:

$$u = A'y.$$

Es gilt dann  $u \sim N_{N \cdot T - m}(0, A'\Sigma A)$  und die Dichte von  $u$  sieht folgendermaßen aus

$$f_u(A'y|\theta) = (2\pi)^{-\frac{N \cdot T - m}{2}} |A'\Sigma A|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}u'[A'\Sigma A]^{-1}u\right\}.$$

Weil  $u$  und  $\hat{\delta}$  stochastisch unabhängig sind, folgt

$$\begin{aligned}
 f_u(A'y|\theta) &= \int_{\mathbb{R}^m} f_{u\hat{\delta}}(A'y, G'y|\theta, \delta) d\delta \\
 &= \int_{\mathbb{R}^m} f_u(A'y|\theta) f_{\hat{\delta}}(G'y|\theta, \delta) d\delta \\
 &= |A, G|^{-1} \int_{\mathbb{R}^m} f_y(y|\theta, \delta) d\delta.
 \end{aligned}$$

Wobei gilt  $G = [X'\Sigma^{-1}X]^{-1} X'\Sigma^{-1}$ . Nutzen wir die beiden folgenden Eigenschaften aus Harville (1974):

$$\begin{aligned}
 |A, G| &= (|(A, G)'(A, G)|)^{\frac{1}{2}} \\
 &= |I|^{\frac{1}{2}} |GG - G'AI^{-1}A'G|^{-\frac{1}{2}} \\
 &= |X'X|^{-\frac{1}{2}}
 \end{aligned}$$

## 2 Statistische Grundlage der Panel-Regression

und

$$(y - X\delta)' \Sigma^{-1} (y - X\delta) = (y - X\hat{\delta})' \Sigma^{-1} (y - X\hat{\delta}) + (\delta - \hat{\delta})' (X' \Sigma^{-1} X) (\delta - \hat{\delta}),$$

folgt

$$f_u(A'y|\theta) = (2\pi)^{-\frac{N \cdot T - m}{2}} |X'X|^{\frac{1}{2}} |\Sigma|^{-\frac{1}{2}} |X' \Sigma^{-1} X|^{-\frac{N \cdot T - m}{2}} \exp(y - X\hat{\delta})' \Sigma^{-1} (y - X\hat{\delta}).$$

Weil gilt

$$\begin{aligned} (y - X\hat{\delta})' \Sigma^{-1} (y - X\hat{\delta}) &= y'(I - GX') \Sigma^{-1} (I - XG') y \\ &= y' \Pi(\theta) y - y' \Sigma^{-1} X [X' \Sigma^{-1} X]^{-1} X' \Sigma^{-1} y + \\ &\quad + y' \Sigma^{-1} X [X' \Sigma^{-1} X]^{-1} X' \Sigma^{-1} X [X' \Sigma^{-1} X]^{-1} X' \Sigma^{-1} y \\ &= y' \Pi(\theta) y, \end{aligned}$$

wobei  $\Pi(\theta)$  in 13 definiert ist, erhält man

$$L_u(\theta) := f_u(A'y|\theta) = (2\pi)^{-\frac{N \cdot T - m}{2}} |X'X|^{\frac{1}{2}} |\Sigma|^{-\frac{1}{2}} |X' \Sigma^{-1} X|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} y' \Pi(\theta) y\right\}. \quad (18)$$

Diese Dichtefunktion nennt man auch REML-Funktion. Logarithmiert man die REML-Funktion (18), erhält man:

$$\begin{aligned} l_u(\theta) &= -\frac{1}{2} [(N \cdot T - m) \ln(2\pi) - \ln |X'X| + \ln |\Sigma|] \\ &\quad - \frac{1}{2} \left[ \ln |X' \Sigma^{-1} X| + y' \Pi(\theta) y \right]. \end{aligned} \quad (19)$$

**Satz 2.16.** Sei  $y^*$  wie im Modell (7) definiert. Setzt man

$$N^* = I_{N \cdot T} - [I_N \otimes C] X (X' [I_N \otimes V(\phi)] X)^{-1} X' [I_N \otimes C'],$$

existiert der REML-Schätzer genau dann, wenn gilt  $N^* y^* \notin \langle Z^* \rangle$ .

**Beweis.** Demidenko und Massam (1999) Satz 3.4 □

**Bemerkung 2.17.** Es ist nicht schwer zu zeigen, dass gilt:

$$N^* y^* \notin \langle Z^* \rangle \iff N^* y \notin \langle Z \rangle.$$

Allerdings hilft uns diese Eigenschaft nicht weiter, weil man das initiale Modell nicht prüfen kann. Aber da die Existenz des ML-Schätzers die Existenz des REML-Schätzers immer impliziert (vergleiche Demidenko und Massam (1999)), können wir nichtsdestotrotz auf eine hinreichende Bedingung der Existenz vom REML-Schätzer schließen.

## 2 Statistische Grundlage der Panel-Regression

**Lemma 2.18.** Den REML-Schätzer für die Varianz  $\Sigma$  erhält man, indem man die folgende Gleichung löst:

$$\text{spur} \left( \Pi(\theta) \frac{\partial \Sigma}{\partial \theta_i} \right) = -y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y, \quad i = 1, \dots, p + 2.$$

*Beweis.* Leitet man die logarithmierte REML-Funktion mithilfe der Rechenregel (14) und (15) ab und setzt sie gleich null, erhält man

$$\text{spur} \left( \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \right) + \text{spur} \left( [X' \Sigma^{-1} X]^{-1} \frac{\partial [X' \Sigma^{-1} X]}{\partial \theta_i} \right) + y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y = 0$$

$$i = 1, \dots, p + 2. \quad (20)$$

Weil gilt

$$\begin{aligned} \text{spur} \left( [X' \Sigma^{-1} X]^{-1} \frac{\partial [X' \Sigma^{-1} X]}{\partial \theta_i} \right) &= -\text{spur} \left( [X' \Sigma^{-1} X] X' \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \Sigma^{-1} X \right) \\ &= -\text{spur} \left( \Sigma^{-1} X [X' \Sigma^{-1} X] X' \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta_i} \right), \end{aligned}$$

lässt sich die Gleichung (20) wie folgt umschreiben

$$\text{spur} \left( \Pi(\theta) \frac{\partial \Sigma}{\partial \theta_i} \right) + y' \frac{\partial \Pi(\theta)}{\partial \theta_i} y = 0, \quad i = 1, \dots, p + 2.$$

□

### 2.1.2 Eigenschaften von ML- und REML-Schätzern

**Korollar 2.19.** Der ML-Schätzer  $\hat{\theta}_{ML}$  ist nur unverzerrt, wenn gilt:

$$\Pi(\theta) = \Sigma^{-1}.$$

Da in der Regel in einem ML-Ansatz die Voraussetzung aus dem Korollar 2.19 eher selten zutrifft, ist der resultierende Schätzer der Varianzkomponenten aufgrund der Unberücksichtigung des Verlusts an Freiheitsgraden – durch die Schätzung von  $\delta$  – verzerrt (vergleiche auch Lancaster (2000)).

Obwohl  $\hat{\theta}_{ML}$  unter angemessenen Regularitätsbedingungen (vergleiche Sweeting, 1980) asymptotisch unverzerrt ist, liegt seine Schwäche immer noch in der finite-sample-Schätzung.

**Korollar 2.20.** Der REML-Schätzer  $\hat{\theta}_{REML}$  ist unverzerrt.



## 2 Statistische Grundlage der Panel-Regression

Im Gegensatz zu  $\hat{\theta}_{ML}$  haben die asymptotischen Eigenschaften des REML-Schätzers für  $\hat{\theta}_{REML}$  nur wenig Aufmerksamkeit in der Literatur gewonnen. Mardia und Marshall (1984) untersuchten die Konsistenz und asymptotischen Eigenschaften von  $\hat{\theta}_{ML}$  in der linearen Regression und gaben die Ergebnisse für gemischte lineare Modelle an. Sweeting (1980) ging anders vor: Er zeigte die Konsistenz und asymptotischen Eigenschaften, indem er auf die Theorie der Funktionalanalysis zurückgriff. Erst später haben sich einige Autoren für die Konsistenz und die asymptotischen Eigenschaften des REML-Schätzers für die Varianz-Parameter interessiert. Cressie und Soumendra (1993) untersuchten diese Eigenschaften unter Verwendung eleganter Argumente aus Sweeting (1980). Alle diese genannten Ergebnisse basieren auf der Annahme, dass  $y$  gemäß  $N(X\delta, \Sigma)$  verteilt ist. Andere Autoren wie Huggins (1993) und Richardson und Welsch (1994) setzen keine Normalverteilung voraus. Die Ergebnisse des ersten Autors basieren auf der Martingaltheorie, nehmen aber eine symmetrische sphärische Verteilung an. Die anderen zeigen sogar, dass auch ohne Verteilungsannahme der REML-Schätzer konsistent sein kann.

Wir betrachten den messbaren Raum  $(\Omega, \mathfrak{F})$ . Sei  $P_\theta$  ein von  $\theta \in \Theta \subset \mathbb{R}^{p+2}$  abhängiges Wahrscheinlichkeitsmaß auf  $(\Omega, \mathfrak{F})$ . Es wird angenommen, dass  $P_\theta$  – mit zugehöriger Dichte  $L_u(\cdot)$  aus (18) – für alle  $\theta \in \Theta$  absolut stetig im Bezug auf das Lebesgue-Maß ist. Sei  $\ell := N \cdot T$  die Anzahl vorliegender Beobachtungen. Da  $\Sigma$  zweimal stetig differenzierbar in  $\theta \in \Theta^\circ$  (Innere von  $\Theta$ ) ist, ist  $l_u(\cdot)$  aus (19) auch zweimal stetig differenzierbar in  $\theta \in \Theta^\circ$ . Die Matrix  $\mathcal{J}_\ell(\theta)$  der partiellen Ableitungen der zweiten Ordnung von  $l_u(\cdot)$  – auf Basis von  $\ell$  Beobachtungen – sieht wie folgt aus (vergleiche Cressie und Soumendra (1993))

$$[\mathcal{J}_\ell(\theta)]_{ij} = \frac{1}{2} (\text{spur}(\Pi(\theta)[\Sigma_{ij} - \Sigma_i \Pi(\theta) \Sigma_j]) - y' \Pi(\theta) [\Sigma_{ij} - 2\Sigma_i \Pi(\theta) \Sigma_j] \Pi(\theta) y)$$

$$i, j = 1, \dots, p+2,$$

wobei gilt  $\Sigma_i = \frac{\partial \Sigma}{\partial \theta_i}$  und  $\Sigma_{ij} = \frac{\partial^2 \Sigma}{\partial \theta_i \partial \theta_j}$ . Sei  $M = [\theta_1^0, \theta_2^0, \dots, \theta_{p+2}^0]$  ( $\theta_i^0 \in \Theta^\circ$ ) eine  $(p+2) \times (p+2)$ -Matrix.  $\mathcal{J}_\ell(M)$  bezeichnet die Matrix mit  $[\mathcal{J}_\ell(M)]_{ij} = \left( \frac{\partial^2 l_u(\theta)}{\partial \theta_i \partial \theta_j} \right) \Big|_{\theta=\theta_i^0}$ ,  $i, j = 1 \dots p+2$ . Für eine beliebige  $l \times r$ -Matrix  $B$  definieren wir die folgende Norme

$$\|B\| = \text{spur}(B'B)^{\frac{1}{2}}.$$

**Definition 2.21.** Wir betrachten den metrischen Raum  $(\mathbb{R}^l, d)$ . Eine Abbildung  $f: \mathbb{R}^l \rightarrow \mathbb{R}$  heißt gleichmäßig stetig genau dann, wenn gilt

$$\forall \varepsilon > 0, \exists \eta > 0 : \forall x_0, x_1 \in \mathbb{R}^l : d(x_1, x_0) < \eta \implies |f(x_1) - f(x_0)| < \varepsilon.$$

## 2 Statistische Grundlage der Panel-Regression

**Definition 2.22.** Gegeben sei eine Funktionenfolge  $(f_\ell)_{\ell \in \mathbb{N}}$  und  $f$ , sodass  $f, f_\ell : \mathbb{R}^l \rightarrow \mathbb{R}$  gilt. Die Folge  $(f_\ell)_{\ell \in \mathbb{N}}$  konvergiert genau dann gleichmäßig gegen  $f$ , wenn gilt

$$\lim_{\ell \rightarrow \infty} \sup_{x \in \mathbb{R}^l} |f_\ell(x) - f(x)| = 0.$$

**Definition 2.23.** Gegeben sei  $\{V, V_\ell, \ell \in \mathbb{N}\} \subset \mathbb{R}^l$  eine Folge von Zufallsvariablen. Die Folge  $(V_\ell)_{\ell \in \mathbb{N}}$  konvergiert schwach gegen  $V$ , wenn für alle beschränkten gleichmäßig stetigen Funktionen  $f : \mathbb{R}^l \rightarrow \mathbb{R}$  die Folge  $(E_\theta f(V_\ell))_{\ell \in \mathbb{N}}$  gleichmäßig gegen  $E_\theta f(V)$  konvergiert.

**Notation 2.24.** Die folgende Bezeichnungen werden künftig adoptiert:

**N.1**  $\rightarrow_u$  steht für eine gleichmäßige Konvergenz in einer kompakten Teilmenge von  $\Theta$ .

**N.2**  $V_\ell \Rightarrow_u V$ , wenn gilt  $E_\theta f(V_\ell) \rightarrow_u E_\theta(f(V))$ .

**Bemerkung 2.25.** (vergleiche Cressie und Soumendra (1993))

$$(V_\ell \Rightarrow_u V) \implies (V_\ell \rightarrow_d V)$$

Wobei  $\rightarrow_d$  für die Konvergenz in Verteilung von Zufallsvariablen steht.

**Annahme (A.1):**

(a)  $\Sigma$  ist zweimal stetig differenzierbar in  $\theta \in \Theta^\circ$ .

(b) Es existieren feste  $(p+2) \times (p+2)$ -Matrizen  $W(\theta)$  und  $\{B_\ell(\theta) : \ell \geq 1\}$  stetig in  $\theta \in \Theta^\circ$ , sodass gilt

$$\|B_\ell(\theta)^{-1}\| \rightarrow_u 0 \text{ und } B_\ell(\theta)^{-1} \mathcal{J}_\ell(\theta) (B_\ell(\theta)^{-1})' \Rightarrow_u W(\theta),$$

wobei  $W(\theta)$  positiv definit für alle  $\theta$  in  $\Theta^\circ$  ist.

(c)  $\forall c > 0, \forall \eta > 0$ ,

1.

$$\sup_{\{\theta, \theta^0 \in \Theta^\circ: \|B_\ell(\theta)'(\theta - \theta^0)\| \leq c\}} \|B_\ell(\theta)^{-1} B_\ell(\theta^0) - I_{p+2}\| \rightarrow_u 0$$

2. für  $M = [\theta_1^0, \theta_2^0, \dots, \theta_{p+2}^0]$ ,

$$A_\ell := \sup_{\{\theta_i^0 (1 \leq i \leq p+2): \|B_\ell(\theta)'(\theta - \theta_i^0)\| < c\}} \|B_\ell(\theta)^{-1} [\mathcal{J}_\ell(M) - \mathcal{J}_\ell(\theta)] (B_\ell(\theta)^{-1})'\|$$

$$P_\theta(A_\ell > \eta) \rightarrow_u 0$$

## 2 Statistische Grundlage der Panel-Regression

**Satz 2.26.** Sei  $\hat{\theta}_\ell$  der auf die  $\ell$  ersten Beobachtungen basierende REML-Schätzer. Dann gilt unter (A.1):

$$B_\ell(\theta)'(\hat{\theta}_\ell - \theta) \implies_u N_{p+2}(0, W(\theta)^{-1})$$

*Beweis.* Sweeting (1980) Satz 1 und 2. □

Bedingung (a) und (c)<sub>2</sub> verlangen die Glattheit von  $\Sigma$  in  $\theta$  und können direkt im Modell geprüft werden. (c)<sub>1</sub> ist abhängig von der Wahl der Matrizen  $\{B_\ell, \ell \geq 1\}$ . Sie können mit einer genügenden Glattheit gewählt werden, um diese Bedingung zu garantieren (vergleiche Cressie und Soumendra (1993)). Folglich werden wir uns auf die Annahme (b) konzentrieren. Für manche Varianzkomponenten kann die Prüfung der Annahme (b) extrem schwierig sein. In solchen Fällen sind die beiden folgenden Bedingungen hinreichend, um (b) zu garantieren.

**Annahme (A.3):** Es gilt

(d) für  $\ell \rightarrow \infty$ ,

$$\sum_{i=1}^{p+2} \sum_{j=1}^{p+2} \sum_{l=1}^{p+2} \sum_{r=1}^{p+2} c_{il}c_{jr} \{ \text{spur}(\Pi(\theta)\Sigma_{lj}\Pi(\theta)\Sigma_{ri}) + \sum_{s=1}^{\ell} (\Sigma^{\frac{1}{2}}\Pi(\theta)\Sigma_{lj}\Pi(\theta)\Sigma^{\frac{1}{2}})_{ss} (\Sigma^{\frac{1}{2}}\Pi(\theta)\Sigma_{ri}\Pi(\theta)\Sigma^{\frac{1}{2}})_{ss} \} \longrightarrow_u 0,$$

(e) für  $\ell \rightarrow \infty$ ,

$$\sum_{i=1}^{p+2} \sum_{i=j}^{p+2} \sum_{l=1}^{p+2} \sum_{r=1}^{p+2} c_{il}c_{jr} \{ \text{spur}(\Pi(\theta)\Sigma_l\Pi(\theta)\Sigma_j\Pi(\theta)\Sigma_r\Pi(\theta)\Sigma_i) + \sum_{s=1}^{\ell} (\Sigma^{\frac{1}{2}}\Pi(\theta)\Sigma_l\Pi(\theta)\Sigma_j\Pi(\theta)\Sigma^{\frac{1}{2}})_{ss} (\Sigma^{\frac{1}{2}}\Pi(\theta)\Sigma_r\Pi(\theta)\Sigma_i\Pi(\theta)\Sigma^{\frac{1}{2}})_{ss} \} \longrightarrow_u 0,$$

wobei gilt  $C_\ell(\theta) = (B_\ell(\theta))^2$  mit  $c_{ij} = (C_\ell(\theta))_{ij}$ ,  $1 \leq i, j \leq p+2$ .

**Annahme (A.4):**

Es existieren nichtstochastische positiv definite Matrizen  $W(\theta)$  und  $\{B_\ell(\theta)\}_{n \geq 1}$ , stetig in  $\theta$ , sodass gilt  $\|B_\ell(\theta)^{-1}\| \longrightarrow_u 0$  sowie  $B_\ell(\theta)^{-1} E_\theta(\mathcal{J}_\ell(\theta)) B_\ell(\theta)^{-1} \longrightarrow_u W(\theta)$ .

**Satz 2.27.** Unter den Annahmen (A.1), (A.2), (A.3) und (A.4) erfüllt  $\hat{\theta}_\ell = \hat{\theta}_{REML}$

$$B_\ell(\theta)'(\hat{\theta}_\ell - \theta) \implies_u N_{p+2}(0, W(\theta)^{-1}).$$

*Beweis.* Cressie und Soumendra (1993). □

**Korollar 2.28.** Unter den Annahmen (A.1), (A.2), (A.3) und (A.4) gilt

$$(E_\theta[\mathcal{J}_\ell(\theta)])^{\frac{1}{2}}(\hat{\theta}_\ell - \theta) \implies_u N_{p+2}(0, I).$$

## 2.2 Bayesianischer Ansatz

Die Anzahl der zur Verfügung stehenden Computer-Pakete für die Optimierung der Maximum-Likelihood-Funktion im Kontext der Panel-Regression wird immer höher (SAS beispielweise Proc Mixed, R-Funktion lme). Obwohl diese Pakete hochentwickelte numerische Verfahren zur Schätzung von festen sowie zufälligen Effekten ausnutzen, sind sie im Allgemeinen weder in der Lage die Varianzinflation aufzunehmen noch die aus der Festlegung der a-priori-Verteilungen entstehenden Effekte einzubeziehen. Die bayesianische Analyse scheint deshalb eine gute Alternative zu sein. Hsiao u.a. (1997) sind davon überzeugt und meinen sogar, dass

*„There is no excuse any longer for not considering Bayesian methods because of practical difficulties. Since the comparison of alternative estimators in this study is undertaken to a large extent using the sampling properties of estimators, econometrician who adopt the sampling theory frame work should find the Bayesian approach more attractive in the light of our finding.“*

Allerdings gibt es oft nicht genügend Informationen über die Modellparameter und sogar sehr oft fehlen a-priori-Verteilungen. Nichtinformativ a-priori-Verteilungen sind deshalb von großem Interesse, da sie keinen Einfluss auf die a posteriori-Verteilung haben und lassen somit die Daten selbst sprechen. Leider ist es sehr schwierig gewöhnliche nichtinformativ a-priori-Verteilungen wie Jeffreys-, Berger-Bernado- oder Matching a-priori-Verteilung herzuleiten. Eine Übersicht über nichtinformativ a-priori-Verteilungen findet man in Kass und Casella (1996). Einfache nichtinformativ a-priori-Verteilungen – beispielweise die Gleichverteilung über die reellen Zahlen – sind in diesem Fall attraktiver, haben aber den Nachteil – dass ohne geeignete Vorsichtsmaßnahme – sie zur Nichtkonvergenz des Gibbs-sampling Algorithmus führen können (vergleiche Hobert und Casella (1996) sowie Sun u.a. (2001)). Lancaster (2004, S. 34) stellte einwandfrei fest, dass a-priori-Verteilungen, die über reelle Zahlen gleichverteilt sind, in der Praxis denjenigen äquivalent sind, die dort – wo die Likelihood nicht vernachlässigbar ist – gleichverteilt sind. Wie sich die a-priori-Verteilung ausserhalb dieses Bereiches verhält, ist absolut unbedeutend für die a-posteriori-Verteilung. Auf dieser Basis können informativ a-priori-Verteilungen in Frage kommen, da sie für einige Verteilungsparameter in einem bestimmten Bereich gleichverteilt sein können. Wir beginnen das hierarchische Modell mit den

folgenden Annahmen

$$\left\{ \begin{array}{l} y|\xi_1, \dots, \xi_p, \delta, \sigma_0^2 \sim N_{N \cdot T}(X\delta + \sum_{i=1}^p Z_i \xi_i, \sigma_0^2 I), \\ \xi_i = (\xi_{i1}, \dots, \xi_{in_i}), \quad i = 1, \dots, p, \\ \xi_{i1}, \dots, \xi_{in_i} | G \sim^{uiv} G, \\ G | \nu_i \sim F_G(\nu_i), \quad \nu_i \sim p(\nu_i), \\ \nu := (\nu_1, \dots, \nu_p) \sim p_\nu(\nu) = \prod_i^p p(\nu_i), \\ \delta \sim p_\delta(\delta), \quad \sigma_0^2 \sim p_0(\sigma_0^2). \end{array} \right. \quad (21)$$

Wobei

$n_i, i = 1, \dots, p$  die Anzahl unabhängiger zufälliger Effekte in  $\xi_i$  ist. Damit gilt  $k_i = \dim(\xi_i) = n_i \cdot \dim(\xi_{i1})$ .  $F_G$  – mit Verteilungsparameter  $\nu$  und Dichte  $f_G$  – ist die Verteilungsfunktion von  $G$ .  $p_\delta, p_0$  sowie  $p_\nu$  sind die a-priori-Verteilungen für  $\delta, \sigma_0^2$  bzw.  $\nu$ .

Im Modell (21) werden die folgenden Annahmen (bedingte Unabhängigkeit) getroffen:

- (f)  $y|\xi_1, \dots, \xi_p$  ist unabhängig von  $\nu_1, \dots, \nu_p$ .
- (g)  $\xi_i|\nu_i, i = 1, \dots, p$  sind sowohl untereinander als auch von  $\delta$  unabhängig.
- (h)  $\delta, \sigma_0^2$  sowie  $\nu_1, \dots, \nu_p$  sind a-priori-unabhängig.

Die Schätzung der Modellparameter erfolgt in diesem Abschnitt durch den Gibbs-sampling-Algorithmus oder den Metropolis-Algorithmus innerhalb des Gibbs-sampling-Algorithmus (vergleiche Tierney (1994)). Wie der Gibbs-sampling-Algorithmus im Kontext der Panel-Regression funktioniert, ist durch die folgenden Schritte aus Gelfand u.a. (1990) zu sehen.

1. Wähle einen Startwert  $s^0 = (\delta^0, \nu_1^0, \dots, \nu_p^0, \sigma_0^{20})$
2. Für  $i = 1, \dots, p + 2$ , ziehe Zufallszahlen aus der bedingten Dichte

$$p(s_i | s_1^{(l)}, \dots, s_{i-1}^{(l)}, s_{i+1}^{(l-1)}, \dots, s_{p+2}^{(l-1)}, y) \quad (22)$$

3. Erhöhe  $l$  um eins und wiederhole ab 2.  $L$ -Mal ( $L$  ist frei wählbar).

Die aus dem Gibbs-sampling-Algorithmus resultierende marginale Verteilung lässt sich durch das folgende Lemma aus Gelfand u.a. (1990) ermitteln.

## 2 Statistische Grundlage der Panel-Regression

**Lemma 2.29.**

$$\hat{p}_{s_i}^L(s_i) = \frac{\sum_{l=1}^L p(s_i | s_1 = s_1^{(l)}, \dots, s_{i-1} = s_{i-1}^{(l)}, s_{i+1} = s_{i+1}^{(l)}, \dots, s_{p+2} = s_{p+2}^{(l)}, \mathbf{y})}{L} \quad (23)$$

**Beweis.**  $s_{-i} := (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{p+2})$

$$\begin{aligned} p_{s_i}(s_i) &= \int p(s_i | s_{-i}, \mathbf{y}) ds_{-i} \\ &= E_{p(s_{-i})} (p(s_i | s_{-i}, \mathbf{y})) \\ &\approx \frac{\sum_{l=1}^L p(s_i | s_1 = s_1^{(l)}, \dots, s_{i-1} = s_{i-1}^{(l)}, s_{i+1} = s_{i+1}^{(l)}, \dots, s_{p+2} = s_{p+2}^{(l)}, \mathbf{y})}{L}. \end{aligned}$$

□

Mit der Bayes'schen Regel

$$p(\theta | \mathbf{y}) = \frac{p(\mathbf{y} | \theta) p(\theta)}{\int p(\mathbf{y} | \theta) p(\theta) d\theta} \propto p(\mathbf{y} | \theta) p(\theta) \quad (24)$$

können wir die für den Gibbs-sampling-Algorithmus notwendige a-posteriori-Verteilung bestimmen. Unter Berücksichtigung der Annahmen – über die bedingte Unabhängigkeit – (f), (g) und (h) im Modell (21) lässt sich die a-posteriori-Verteilung wie folgt ermitteln (vergleiche Hobert und Casella (1996))

$$p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y}) \propto f(\mathbf{y} | \delta, \xi_1, \dots, \xi_p, \sigma_0^2) \left( \prod_i^p f_G(\xi_i | \nu_i) \right) p_\delta(\delta) p_0(\sigma_0^2) p_\nu(\nu). \quad (25)$$

Somit folgt beispielweise, dass gilt  $p(\sigma_0^2 | \nu, \xi_1, \dots, \xi_p, \delta, \mathbf{y}) = \frac{p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y})}{\int p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y}) d\sigma_0^2}$  und jede bedingte Dichte aus dem Gibbs-Sampling-Algorithmus lässt sich so ähnlich bestimmen.

In dem Bayesianischen Ansatz ist  $p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y})$  die bedingte Dichte der unbekannt Parameter gegeben die Daten. Es ist dann ersichtlich, dass die Unzulässigkeit einer der a-priori-Verteilungen dazu führen kann, dass die a-posteriori-Verteilung  $p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y})$  auch nicht zulässig ist. Dies würde für diesen Fall darauf hindeuten, dass es keine Dichte gibt, die theoretisch der aus dem Gibbs-Sampling-Algorithmus entspricht. Da zum Anderen  $\int p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | \mathbf{y}) d\sigma_0^2$  konstant in Bezug auf  $\sigma_0^2$  ist, ist  $p(\sigma_0^2 | \nu, \xi_1, \dots, \xi_p, \delta, \mathbf{y})$  der rechten Seite von (25) proportional. Daraus könnte man die Schlussfolgerung ziehen, dass die Kenntnis über

die bedingte Dichten aus dem Gibbs-Sampling-Algorithmus die Zulässigkeit von  $p(\sigma_0^2, \nu, \xi_1, \dots, \xi_p, \delta | y)$  voraussetzt. Leider ist das nicht immer der Fall. Gegenbeispiele dafür findet man bei Hobert und Casella (1996). Solche Gegenbeispiele lassen uns eine gefährliche Situation aufgreifen, wo der Gibbs-Sampling-Algorithmus zu einer vernünftigen Inferenz über eine nicht existierende a-posteriori-Verteilung führt. Aus diesem Grund ist es immer notwendig die Zulässigkeit der a-posteriori-Verteilung zu beweisen, wenn eine der a-priori-Verteilungen unzulässig ist.

Nach dem Lauf des Gibbs-sampling-Algorithmus wird eine Kette für jeden Parameter erzeugt, die gegen die marginale Verteilung konvergiert, wenn die Stationarität erreicht ist. Einen Überblick über Ansätze zur Prüfung der Stationarität findet man bei Cowles und Carlin (1996). Für diese Arbeit werden wir uns auf die Gelman-Rubin-Diagnostik aus Brooks und Gelman (1998) beschränken. Diese Konvergenzdiagnostik lässt sich in den folgenden Punkten beschreiben:

- (i) Lass  $K$  Ketten ( $K, 2$ ) mit einer Länge – genügend – größer  $L$  und unterschiedlichen Startwerten laufen
- (ii) Behalte die  $L$  letzten Elemente der Kette
- (iii) Für jeden Parameter  $s_i, i = 1, \dots, p + 2$ , berechne die Varianz innerhalb der Kette  $W = \frac{1}{K} \sum_{k=1}^K \left[ \frac{1}{L-1} \sum_{l=1}^L (s_{ikl} - \bar{s}_{ik})^2 \right]$  sowie die Varianz zwischen den Ketten  $B = \frac{L}{K-1} \sum_{k=1}^K (\bar{s}_{ik} - \bar{\bar{s}}_i)^2$  ( $\bar{\bar{s}}_i = \frac{1}{LK} \sum_{k=1}^K \sum_{l=1}^L s_{ikl}$ )
- (iv) Berechne die Parametervarianz als  $\hat{\sigma}_{s_i}^2 = (1 - \frac{1}{L})W + \frac{1}{L}B$ . Sei  $\hat{V} = \hat{\sigma}_{s_i}^2 + \frac{B}{LK}$
- (v) Berechne den Potenzial-Reduktion-Faktor  $\hat{R} = \frac{\hat{V}}{\hat{\sigma}_{s_i}^2}$ . Wenn gilt  $\hat{R} \approx 1$ , dann konvergieren jede der generierten Ketten. Ansonsten lasse die Kette mit einer grösseren Länge laufen, um die Kovergenz zu erhöhen.

### 2.2.1 Normalverteilung für zufällige Effekte

Für  $G$  kommt sehr oft in der Literatur die Normalverteilung in Frage. Einen Überblick dafür findet man im Lin und Zhang (1999), Fahrmeir und Lang (2001), Ruppert u.a. (2003) sowie Durban u.a. (2005). Eine Besonderheit des Spezialfalles  $\xi_i \sim N_{k_i}(0, \sigma_i^2 \mathbf{I})$ ,  $i = 1, \dots, p$  – das heißt,  $G | v_i = \sigma_i^2 \sim N_{\dim(\xi_{i1})}(0, \sigma_i^2 \mathbf{I})$  ist, dass die a-posteriori-Verteilung im Model (21) – unter der Nichtinformativität der a-priori-Verteilungen – mit der REML-Funktion eine enge Beziehung hat, die sich im folgenden Satz festhalten lässt.

## 2 Statistische Grundlage der Panel-Regression

**Satz 2.30.** Gelten  $G|v_i = \sigma_i^2 \sim N_{k_i}(0, \sigma_i^2 I)$  mit  $p_i(\sigma_i^2) \propto 1$ ,  $i = 1, \dots, p$  sowie  $p_\delta(\delta) \propto 1$  und  $p_0(\sigma_0^2) \propto 1$ , so folgt

$$\begin{aligned} p(\sigma_0^2, \dots, \sigma_p^2 | y) &= \int_{\mathbb{R}^{k_1}} \dots \int_{\mathbb{R}^{k_p}} \int_{\mathbb{R}^m} p(\sigma_0^2, \dots, \sigma_p^2, \xi_1, \dots, \xi_p, \delta | y) d\delta d\xi_1 \dots d\xi_p \\ &= \frac{|X' \Sigma^{-1} X|^{-\frac{1}{2}}}{(2\pi)^{\frac{N \cdot T - m}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} y' \Pi(\theta) y\right\}. \end{aligned} \quad (26)$$

Wobei  $\Sigma = \sum_{i=1}^p \sigma_i^2 Z_i Z_i' + \sigma_0^2 I$ .

*Beweis.* Searle u.a. (1992, S. 322–325). □

**Satz 2.31.** Sind die Voraussetzungen vom Satz 2.30 erfüllt, dann ist die a-posteriori-Verteilung von  $(\delta, \xi_1, \dots, \xi_p, \sigma_0^2, \dots, \sigma_p^2)$  zulässig.

*Beweis.* Sun u.a. (2001). □

Für die Implementierung des Gibbs-sampling-Algorithmus – bei normalverteilten zufälligen Effekten – werden wir folgende Annahmen (schwache Informativität der a-priori-Verteilungen) über die Parameter treffen. Es wird angenommen, dass  $\delta | \mu_\delta, \Sigma_\delta$  gemäß  $N_m(\mu_\delta, \Sigma_\delta)$  verteilt ist. Es ist üblich für  $\Sigma_\delta$  die Struktur einer Diagonalmatrix mit großen Zahlen anzunehmen, was in der Praxis einer nichtinformativen a-priori-Verteilung für  $\delta$  entspricht (vergleiche Lancaster (2004, S. 34) sowie Zhao u.a. (2006)). Wir wählen  $\mu_\delta = 0_m$  sowie  $\Sigma_\delta = 10.000 I_m$ . Da die Unabhängigkeit der Modellparameter vorausgesetzt ist, lässt sich die Varianz der zufälligen Effekte als Summe der einzelnen Varianzkomponenten ( $\sum_{i=1}^p \sigma_i^2 Z_i Z_i'$ ) geben. Hierfür kommt die inverse Gammaverteilung (IG) als a-priori-Verteilung für  $\sigma_i^2$  sehr oft in der Literatur in Frage (vergleiche Gelfand u.a. (1990), Hobert und Casella (1996) sowie Heinzl u.a. (2012)). Wir nehmen dann an, dass  $\sigma_i^2$ ,  $i = 1, \dots, p$  – wie auch bei Heinzl u.a. (2012) – gemäß  $IG(0,005; 0,005)$  verteilt sind und machen die gleiche Annahme für  $\sigma_0^2$ .

Da die REML-Schätzer für die Varianzparameter  $(\sigma_0^2, \dots, \sigma_p^2)$  durch Anwendung des Newton-Raphson-Algorithmus auf die a-posteriori-Verteilung  $p(\sigma_0^2, \dots, \sigma_p^2 | y)$  in (26) zu bestimmen sind, ist es ersichtlich, dass der REML-Ansatz ein sehr guter Kompromiss zwischen Frequentistischen und Bayesianischen Methoden ist, da sich beide Methoden für normalverteilte zufällige Effekte nur noch im Algorithmus unterscheiden. Obwohl die Wahl der Normalverteilung für  $G|v$  mathematisch vorteilhaft ist, kann es in der Praxis aufgrund der Symmetrie sowie Unimodalität bedenklich sein. Weil die Annahme über die Verteilung für unbeobachtbare ökonomische Faktoren ist, ist es schwierig diese Eigenschaften auf Basis einer Schätzung zu bestätigen. Mögliche Schiefe und Multimodalität können unersichtlich



bei der Prüfung der Normalverteilung im Bezug auf die Schätzung der zufälligen Effekte sein. Zhang und Davidian (2001) stellten ein Histogramm geschätzter individuenspezifischen Effekte aus einzelnen Kleinste-Quadrate-Schätzungen von Framingham-Cholesterol-Daten dar, das die Unangemessenheit normalverteilter zufälliger Effekte deutlich zeigt. Wir wollen eine Klasse flexibler Verteilungen für zufällige Effekte verwenden, damit Einflüsse von restriktiven parametrischen Modellen auf Schlussfolgerungen vermieden werden können. Heinzl u.a. (2012) wandten die Panel-Regression auf die Body-Mass-Index-Daten der LISA-Studie an und fanden, dass das Modell mit der Dirichlet-Prozess-Mischung (DPM) (vergleiche Ferguson (1973)) für zufällige Individueneffekte deutlich besser als der REML-Ansatz im Sinn der Residuen-Quadratsumme aus zufälligen Effekten ist. Die Autoren schlagen deshalb für  $G$  die DPM als Alternative vor. Jara u.a. (2009) sprechen mehr für Polya-Baum-Mischungen (PBM). DPM sowie PBM haben den Vorteil, dass sie die Abweichung von der Normalverteilung in Betracht ziehen können. Im Allgemeinen führen diese beide a-priori-Verteilungen zum Klumpeneffekt. Dies ist gerade in unserem Fall von großem Interesse, weil der von dem Stichprobenmuster abweichende Effekt identifiziert werden soll. Aufgrund der höheren Flexibilität vom PBM gegenüber DPM (vergleiche Hanson (2006)) legen wir großen Wert auf die PBM a-priori-Verteilung für die vorliegende Arbeit.

### 2.2.2 Polya-Baum für zufällige Effekte

Polya-Bäume sind konjugierte Verteilungen für zufällige Wahrscheinlichkeitsmaße, die als Verallgemeinerung des Dirichlet-Prozesses betrachtet werden können (vergleiche Ferguson (1974)). Lavine (1992), Lavine (1994) sowie Paddock u.a. (2003) stellten die Theorie des Polya-Baums vor und erläuterten sie anhand von Daten. Hanson und Johnson (2002) modellierten die Regressionsfehler mit dem Polya-Baum. Paddock (2002), Hanson (2006) sowie Jara u.a. (2009) verallgemeinerten univariate Polya-Bäume zu multivariaten Polya-Bäumen.

Wir betrachten den separablen messbaren Raum  $(\Omega, \mathfrak{F})$ . Seien  $E = \{0,1\}$  und  $E^m = \prod_0^m E$ ,  $m \in \mathbb{N}_0$ . Dabei ist  $E^0 = \emptyset$ . Sei außerdem  $\pi_0 = \{\Omega\}$ ; für  $j \in \mathbb{N}$  definieren wir  $\pi_j := \{B_{\epsilon_{11}\dots\epsilon_{1j}\dots;\epsilon_{q1}\dots\epsilon_{qj}} : \epsilon_{ij} \in E\}$  als Partition von  $\Omega$ .  $\pi_{j+1}$  lässt sich als Verfeinerung von Mengen aus  $\pi_j$  unter  $2^q$  Mengen erhalten. Für den Fall  $\Omega = \mathbb{R}^2$  ist die Ebene in vier Mengen partitioniert.  $\pi_1$  besteht aus den vier Mengen. Dann ist jede dieser Menge weiter in vier Mengen partitioniert. Die daraus resultierenden Mengen – insgesamt 16 – bilden  $\pi_2$ . Weitere Partitionierungen  $\pi_j$ ,  $j > 3$  lassen sich so ähnlich bilden.

**Definition 2.32.** Sei  $\Pi = \{\pi_j, j \in \mathbb{N}_0\}$  eine Menge von Partitionen von  $\Omega$  – wie oben

## 2 Statistische Grundlage der Panel-Regression

beschrieben – sowie  $\mathcal{A} = \{\mathbf{a}_{\epsilon_1 \dots \epsilon_q} : \epsilon_1 \dots \epsilon_q \in E^{j-1}, j \in \mathbb{N}\}$  eine Menge nichtnegativer Vektoren. Eine  $q$ -variates zufälliges Wahrscheinlichkeitsmaß  $G$  auf  $\Omega$  hat eine multivariate Polya-Baum-Verteilung mit Parameter  $(\Pi, \mathcal{A})$ , wenn eine Familie  $\mathcal{Y} = \{\mathbf{Y}_{\epsilon_1 \dots \epsilon_q} : \epsilon_1 \dots \epsilon_q \in E^{j-1}, j \in \mathbb{N}\}$  von Zufallsvektoren existiert, sodass die folgenden Eigenschaften gelten:

- (i) Alle Zufallsvektoren in  $\mathcal{Y}$  sind unabhängig von einander.
- (ii) Für  $j \in \mathbb{N}$  und für jede  $\epsilon_1, \dots, \epsilon_q \in E^{j-1}$ ,  $\mathbf{Y}_{\epsilon_1; \dots; \epsilon_q} \sim \text{Dirichlet}(\mathbf{a}_{\epsilon_1; \dots; \epsilon_q})$ , wobei  $\mathbf{Y}_{\epsilon_1; \dots; \epsilon_q} = \{Y_{\epsilon_1 \epsilon_1 \dots \epsilon_q \epsilon_q} : \epsilon_1, \dots, \epsilon_q \in E\}$  und  $\mathbf{a}_{\epsilon_1; \dots; \epsilon_q} = \{a_{\epsilon_1 \epsilon_1 \dots; \epsilon_q \epsilon_q} : \epsilon_1, \dots, \epsilon_q \in E\}$ .
- (iii) Das  $q$ -variates zufällige Wahrscheinlichkeitsmaß  $G$  ist mit  $\mathcal{Y}$  durch die folgende Beziehung verbunden

$$G(B_{\epsilon_{11} \dots \epsilon_{1j}; \dots; \epsilon_{q1} \dots \epsilon_{qj}}) = \prod_{l=1}^j Y_{\epsilon_{11} \dots \epsilon_{1l}; \dots; \epsilon_{q1} \dots \epsilon_{ql}}.$$

Für die Partitionierung stehen zwei Ansätze zur Verfügung (vergleiche Jara u.a. (2009)). Beide Methoden partitionieren den Raum  $\Omega$  für zufällige Effekte, so dass sich die zufälligen Effekte innerhalb jeder Menge in der Partition auf eine stetige Verteilung (Gleichverteilung oder Normalverteilung) ausrichten. Es ergibt sich eine multivariate Polya-Baum-Mischung (MPBM)-Verteilung. Wir nehmen hier an, dass  $G$  innerhalb von  $\pi_j$  gemäß  $N_q(\mu_*, \Sigma_*)$  verteilt ist. Hanson (2006) schlägt hierfür vor, dass man das cartesianische Produkt von Intervallen aus Quantile der univariaten Standardnormalverteilung bildet und es mit der Abbildung  $u \mapsto \mu_* + \Sigma_*^{\frac{1}{2}} u$  transformiert, wobei  $\Sigma_*^{\frac{1}{2}}$  eine Wurzel von  $\Sigma_*$  repräsentiert. Die Mengen aus dem cartesianischen Produkt lassen sich in der  $j$ ten Partition wie folgt bilden

$$B_o(j, k) = \left( \Phi^{-1}\left(\frac{k_1 - 1}{2^j}\right), \Phi^{-1}\left(\frac{k_1}{2^j}\right) \right) \times \dots \times \left( \Phi^{-1}\left(\frac{k_q - 1}{2^j}\right), \Phi^{-1}\left(\frac{k_q}{2^j}\right) \right),$$

$$k = (k_1, \dots, k_q), k_i \in \{1, \dots, 2^j\}, i = 1, \dots, q. \quad (27)$$

Wendet man die genannte Abbildung auf  $B_o(j, k)$  in (27), so erhält man die Menge

$$B(j, k) = \{\mu_* + \Sigma_*^{\frac{1}{2}} u : u \in B_o(j, k)\},$$

bei der  $G$  eine  $q$ -variates Normalverteilung mit Parametern  $(\mu_*, \Sigma_*)$  hat. Daraus kann man die  $j$ te Partition  $\pi_j = \{B(j, k) : k \in \{1, \dots, 2^j\}^q\}$  erzeugen und die Menge

## 2 Statistische Grundlage der Panel-Regression

der Partition  $\Pi^{\mu_* \Sigma_*} = \{\pi_j, j \in \mathbb{N}_0\}$  bilden. In der Regel richten sich die Tiefe der Partitionierung  $\mathbf{J}$  für  $\xi_{i1}, \dots, \xi_{in_i} | G \sim^{iiv} G$  nach der Faustregel  $\mathbf{J} \approx \log_{2^q} n_i$ . Damit fällt ein einziger zufälliger Effekt unter  $B(\mathbf{J}, k)$ .

Sobald die Partition abgeschlossen ist, können wir uns um die positiven Vektoren kümmern. Hanson (2006) betrachtet  $\mathcal{A}^c = \{\mathbf{a}^c(j, r) : r \in \{1, \dots, 2^{j-1}\}^q, j \in \mathbb{N}\}$ , wobei  $\mathbf{a}^c(j, r) = 1_{pq} \varrho(c, j)$ .  $\{\varrho(c, j), j \in \mathbb{N}\}$  bestimmt, inwieweit  $G$  von der Zentrierungsverteilung abweichen kann. Dadurch entsteht die Flexibilität von PBM gegenüber der Normalverteilung. Für  $\varrho(c, j) = 1$  oder  $\varrho(c, j) = j^2$  ist  $G$  stetig. Hanson und Johnson (2002) fanden, dass  $G$  bezüglich des Lebesgue-Maßes – für  $\varrho(c, j) = cj^2$ ,  $c > 0$  – absolut stetig ist. Sehr große Werte von  $c$  führen dazu, dass die Verteilung von  $G$  näher an der Basisverteilung – in diesem Fall  $N_q(\mu_*, \Sigma_*)$  – ist. Für  $c \rightarrow \infty$  ist  $G$  sogar mit Wahrscheinlichkeit 1 gemäß  $N_q(\mu_*, \Sigma_*)$  verteilt. Kleinere Werte von  $c$  führen dagegen dazu, dass  $E(G)$  näher an der empirischen Verteilung der zufälligen Effekte sind. Deshalb liefert das Polya-Baum-Modell – wie auch bei dem Dirichlet-Prozess – eine Mischung aus der empirischen Verteilung und die Ausrichtungsverteilung. Sehr kleine Werte von  $c$  ( $c \rightarrow 0$ ) haben allerdings den Nachteil, dass die erzeugten Zufallszahlen in der Praxis nicht genügend gemischt sind. Angesicht der Auswirkung von  $c$  auf die generierte Kette schlägt Hanson (2006)  $c = 1$  für eine gute Mischung der Kette sowie  $c = 10$  für eine Sensitivitätsanalyse vor.

In der obigen Partionierung ist festzustellen, dass  $\Sigma_*^{\frac{1}{2}}$  die Richtung der Partitionierung bestimmt, da sich die Wurzeln einer Matrix mit ihren Eigenvektoren unterscheiden. Jara u.a. (2009) schlagen deshalb eine Parametrisierung vor, welche den Effekt der gewählten Wurzel auf die Partition mäßigt. Wir betrachten dazu eine beliebige Wurzel  $U$  von  $\Sigma_*$ , das heißt, eine reelle und nichtsinguläre  $q \times q$  Matrix, so dass gilt  $\Sigma_* = U'U$ .  $U$  lässt sich weiter zerlegen, indem das folgende Lemma aus Jara u.a. (2009) genutzt ist.

**Lemma 2.33.** *Für eine reelle nichtsinguläre  $q \times q$  Matrix  $U$  existiert eine orthogonale Matrix  $O$ , sodass gilt  $U = OM$ . Wobei  $M$  eine reelle obere Dreiecksmatrix mit positiven Diagonalelementen ist.*

Sei  $R$  die eindeutige obere Dreiecksmatrix aus der Cholesky-Zerlegung von  $\Sigma_*$ . Für jede Wurzel  $U$  von  $\Sigma_*$  existiert eine orthogonale Matrix  $O$  – aus dem Lemma 2.33 –, sodass gilt  $U = OR$ . Damit sieht die neue Partitionierung folgendermaßen aus

$$B^{\mu_* \Sigma_* O}(j, k) = \{\mu_* + M'O'u : u \in B_o(j, k)\}. \quad (28)$$

Die  $j$ te Partition und damit die Menge aller Partitionen lassen sich auf dieser Basis definieren:  $\pi_j^{\mu_* \Sigma_* O} = \{B^{\mu_* \Sigma_* O}(j, k) : k \in \{1, \dots, 2^j\}^q\}$  und  $\Pi^{\mu_* \Sigma_* O} = \{\pi_j^{\mu_* \Sigma_* O}, j \in \mathbb{N}_0\}$ .

## 2 Statistische Grundlage der Panel-Regression

$\mathbb{N}_0\}$ . Die Cholesky-Zerlegung von  $\Sigma_*$  ( $U = R$ ) tritt auf, wenn gilt  $O = I_q$ . Die symmetrisch und asymmetrisch Wurzelzerlegungen tauchen dagegen auf, wenn gilt  $U = O_E \Lambda^{\frac{1}{2}}$  beziehungsweise  $U = O_E \Lambda^{\frac{1}{2}} O'_E$ , wobei  $O_E$  die Orthogonalmatrix aus orthogonalen Eigenvektoren sowie  $\Lambda$  die Diagonalmatrix aus Eigenwerten von  $\Sigma_*$  ist. Wenn  $O$  in dem Vektorraum  $\mathcal{O}^q$  variiert, so ist zu erwarten, dass jede Wurzel von  $\Sigma_*$  an der Partitionierung teilnimmt, was den Effekt einer bestimmten Wurzel auf die Partitionierung und die damit verbundene a-posteriori Inferenz mindert. Jara u.a. (2009) schlagen deshalb für  $O$  die Gleichverteilung auf  $\mathcal{O}^q$  (Haarverteilung) als a-priori-Verteilung (vergleiche Wijsman (1990) sowie James (1954)) vor.

**Definition 2.34.** Eine Wahrscheinlichkeitsverteilung  $F$  auf  $\mathcal{O}^q$  heißt Haarverteilung, wenn für eine gegebene Matrix  $O \in \mathcal{O}^q$  gilt

$$F(O) = F(AO) = F(OB), \forall A, B \in \mathcal{O}^q. \quad (29)$$

Wir erhalten das folgende Modell für die zufälligen Effekte.

$$\begin{cases} \xi_{i1}, \dots, \xi_{in_i} | G \sim^{uiv} G, \quad i = 1, \dots, p, \\ G | (c_i, \mu_{i*}, \Sigma_{i*}, O_i) \sim MPB(\Pi^{\mu_{i*} \Sigma_{i*} O_i}, \mathcal{A}_i^c), \\ \mu_{i*} | \mu_{\xi_i}, S_{\xi_i} \sim N_q(\mu_{\xi_i}, S_{\xi_i}), \quad \Sigma_{i*} | \nu_{i0}, M_i \sim IW(\nu_{i0}, M_i), \\ O_i \sim Haar(q), \quad c_i | a_{i0}, b_{i0} \sim \Gamma(a_{i0}, b_{i0}). \end{cases} \quad (30)$$

Wobei sich  $\Gamma$  und  $IW$  auf die Gamma- beziehungsweise Inverse-Wishart-Verteilung beziehen.

Von Interesse sind die Parameter der zufälligen Effekte. Im folgenden Satz werden diese Kennzahlen ermittelt.

**Satz 2.35.** Sei  $\mathbf{w} \sim G$  und  $G \sim MPB^J(\Pi^{\mu \Sigma O}, \mathcal{A}^c)$ , wobei  $MPB^J(\Pi^{\mu \Sigma O}, \mathcal{A}^c)$  der teilweise festgelegte multivariate Polya-Baum bis zur Stufe  $\mathbf{J}$  ist. Dann folgt

$$\mu_G = E_G(\mathbf{w}) = \mu + U \left( \sum_{k \in I_q^J} G(B(\mathbf{J}, k)) \mu_{B_0(\mathbf{J}, k)}^1 \right) \quad (31)$$

und

$$\begin{aligned} \Sigma_G &= var_G(\mathbf{w}) \\ &= U \left[ \sum_{k \in I_q^J} G(B(\mathbf{J}, k)) \mu_{B_0(\mathbf{J}, k)}^2 [U^{-1}(\mu_G - \mu)][U^{-1}(\mu_G - \mu)]' \right] U', \end{aligned} \quad (32)$$

## 2 Statistische Grundlage der Panel-Regression

wobei  $I_q^J = \{1, \dots, 2^J\}^q$ ,  $\mu_{B_0(\mathbf{J}, k)}^1 = 2^J (\Phi(\kappa_{11}) - \Phi(\kappa_{12}), \dots, \Phi(\kappa_{q1}) - \Phi(\kappa_{q2}))$  und  $\mu_{B_0(\mathbf{J}, k)}^2$  eine symmetrische Matrix mit  $\mu_{B_0(\mathbf{J}, k)}^2_{ii} = 1 - 2^J [\kappa_{i2} \Phi(\kappa_{i2}) - \kappa_{i1} \Phi(\kappa_{i1})]$  repräsentiert. Desweiteren gilt  $\mu_{B_0(\mathbf{J}, k)}^2_{ij} = 4^J [\Phi(\kappa_{i1}) - \Phi(\kappa_{i2})][\phi(\kappa_{j1}) - \Phi(\kappa_{j2})]$ ,  $j \neq i$ , wobei  $\kappa_{i1} = \Phi^{-1}(\frac{\kappa_i - 1}{2^j})$ ,  $\kappa_{i2} = \Phi^{-1}(\frac{\kappa_i}{2^j})$ ,  $i, j = 1, \dots, q$ .

**Beweis.** Jara u.a. (2009). □

Mit den Parametern aus Modell (30) ergänzen wir die a-posteriori-Verteilung (25) und erhalten:

$$\begin{aligned}
 p(\delta, \xi_{1:p}, \sigma_0^2, c_{1:p}, \mu_{1:p*}, \Sigma_{1:p*}, O_{1:p} | y) &\propto p(y | \delta, \xi_{1:p}, \sigma_0^2) \times \\
 &\left[ \prod_{i=1}^p \prod_{l=1}^{n_i} p(\xi_{i_l} | c_i, \mu_{i*}, \Sigma_{i*}, O_i) \right] \times \\
 &p(\delta) p(\sigma_0^2) \prod_{i=1}^p p(\mu_{i*}) p(O_i) p(\Sigma_{i*}) p(c_i).
 \end{aligned} \tag{33}$$

Der Metropolis-Algorithmus innerhalb des Gibbs-samplings (vergleiche Tierney (1994)) wird bei dieser a-posteriori-Verteilung zur Aktualisierung der Parameter eingesetzt. Dabei werden Zufallszahlen aus den einzelnen bedingten Verteilungen ( $p(\xi_{i_l} | \xi_{i_{-l}}, c_i, \mu_{i*}, \Sigma_{i*}, O_i, \sigma_0^2, y)$ ,  $p(c_{1:p}, \mu_{1:p*}, \Sigma_{1:p*}, O_{1:p} | \xi_{1:p}, \delta, \sigma_0^2, y)$  sowie  $p(\delta | \xi_{1:p}, \sigma_0^2, c_{1:p}, \mu_{1:p*}, \Sigma_{1:p*}, O_{1:p}, y)$ ), wie in Jara u.a. (2009) beschrieben ist, gezogen, wobei gilt  $\xi_{i_{-l}} = (\xi_{i_1}, \dots, \xi_{i_{l-1}}, \xi_{i_{l+1}}, \xi_{i_{n_i}})$ .

In der aktuellen Implementierung im Paket *DPpackage* werden nur Modelle berücksichtigt, die so transformiert werden können, dass im hierarchischen Modell (21) gilt  $p = 1$ . Die im Abschnitt 3.2.4 definierten Modelle besitzen nicht diese Eigenschaft (die transformierten zufälligen Effekte sind nicht stochastisch unabhängig), sodass keine Annahme über die PBM-Verteilung dort getroffen wird.

Die Informationen über a-priori-Verteilungen der festen Effekte sowie der Reststörgrößen sind Abschnitt 2.2.1 zu entnehmen. Damit die Basisverteilung keinen großen Einfluss auf die a-posteriori-Verteilung hat, nehmen wir hier – wie auch bei den festen Effekten – an, dass gilt  $\mu_\xi = 0_q$  sowie  $S_\xi = 10.000 I_q$ . Wir unterstellen  $M = I_q$  und  $\nu_0 = q + 2$ , sodass gilt  $E(\Sigma_*) = I_q$ . Außerdem setzen wir  $a_0 = b_0 = 10$  und erhalten  $E(c) = 1$  sowie  $\text{var}(c) = 0,1$ , was für eine gute Mischung der Kette spricht.

### **3 Korrelation von Ausfallwahrscheinlichkeiten**

Die Ausfallwahrscheinlichkeiten von Schuldnern sind die wesentlichen Komponenten zur Verlustabschätzung in einem Kreditportfolio. Bluhm u.a. (2003) stellten drei Vorgehen zur Quantifizierung dieser Kenngrößen vor:

Das erste basiert auf den historischen Ausfallraten von Ratingagenturen wie Moody's oder Standard & Poor's. Tatsächlich sind ihre Datenbanken nicht vorausschauend und enthalten eher die in Bezug auf die Kreditratings eingestufteten Daten, sodass die wichtigen Informationen über einen bestimmten Schuldner vernachlässigt werden. Außerdem zeigten Lando und Skødeberg (2002) das Problem der Verspätung zwischen Ratingveränderungen und dem Herunterstufen bzw. Höherstufen von Unternehmen auf. Ein solches Phänomen könnte zur falschen Einschätzung der Ausfallwahrscheinlichkeit von Unternehmen führen.

Ausfallwahrscheinlichkeiten können auch durch die auf die Optionstheorie von Black und Cox (1976) sowie Merton (1974) basierenden Modelle bestimmt werden. Gemäß dieser Modelle fällt ein Unternehmen aus, wenn sein Eigenkapitalwert unter den Wert seines Kredites vor oder zum Zeitpunkt der Laufzeit fällt. Diese Methode lässt sich kaum zum Einsatz bringen, da sich das zugrunde liegende Eigenkapital nicht direkt beobachten lässt. Zudem variiert die Kapitalstruktur über die Zeit.

Schlussendlich, das dritte Vorgehen, das in dieser Arbeit behandelt wird: es beruht auf dem Konzept der Bewertung von Derivaten, die das Ausfallrisiko von Schuldnern enthält. Wichtigste Kreditderivate für die Messung von Ausfallwahrscheinlichkeiten sind CDS, da sie sich kaum durch die Illiquidität beeinflussen lassen, die wiederum zur falschen Einschätzung von Ausfallwahrscheinlichkeiten führen kann (vergleiche Longstaff u.a. (2005)).

Hull u.a. (2004) wiesen darauf hin, dass ein positiver Zusammenhang zwischen der CDS-Prämie und dem Kreditrating vorliegt. Ericsson u.a. (2009) bestätigten, dass CDS-Prämien durch Variablen erklärt werden können, die für das Kreditrisiko entscheidend sind. Jorion und Zhang (2006) sowie Cariboni u.a. (2011) empfahlen, Ausfallwahrscheinlichkeiten von Banken aus CDS zu schätzen. Zhang u.a. (2012) untersuchten die bedingten Ausfallwahrscheinlichkeiten von EU-Mitgliedsstaaten mittels CDS-Prämien. Die CDS-Prämie kommt somit als wichtigste Messung für die Quantifizierung von Ausfallwahrscheinlichkeiten in Frage und stellt sogar die wesentliche und relevante Information für das Kreditgeschäft dar (vergleiche Norden und Wagner (2008)). Dieser Zusammenhang wird im folgenden erstellt.

### 3.1 Finanzökonomische Theorie

Betrachtet wird eine Anleihe  $B$  (Bond) mit zugehöriger Ausfallwahrscheinlichkeit  $PD$  (Probability of Default) sowie Verlustausfallquote  $LGD$  (Loss Given Default). Wir folgen dem Ansatz des Basler Ausschuss für Bankenaufsicht (vergleiche für Bankenaufsicht (2004) und Jorion (2010)) und bestimmen den erwarteten Verlust  $E[L]$  (Expected Loss) als

$$E[L] = B \cdot PD \cdot LGD \iff PD = \frac{E[L]}{B \cdot LGD}.$$

In einem effizienten Markt kann man davon ausgehen, dass der erwartete Verlust gerade den Versicherungskosten entspricht, die sich als Produkt ( $sp \cdot B$ ) von CDS-Prämie (spread;  $sp$ ) und Anleihe ( $B$ ) bestimmen lassen. Somit ergibt sich (vergleiche Brigo und Mercurio (2006)):

$$PD = \frac{sp \cdot B}{B \cdot LGD} = \frac{sp}{LGD}.$$

Die Verlustausfallquote wird oft deterministisch modelliert (vergleiche Standard & Poors (2006), Brigo und Chourdakis (2009), Collin-Dufresne (2009) sowie Fitch Ratings (2010)). Bei Standard & Poors (2006) und Fitch Ratings (2010) lag die Verlustausfallquote bei etwa 60% ( $LGD \approx 60\%$ ). Somit hängt die Ausfallwahrscheinlichkeit einer Bank ausschließlich von deren CDS-Prämie  $sp$  ab. Die Ausfallwahrscheinlichkeitskorrelation zwischen zwei beliebigen Banken  $i$  und  $j$  lässt sich dann wie folgt ausdrücken

$$\text{korr}(PD_i, PD_j) = \text{korr}\left(\frac{sp_i}{LGD_i}, \frac{sp_j}{LGD_j}\right) = \text{korr}(sp_i, sp_j).$$

Dies bietet uns die Möglichkeit, Ausfallwahrscheinlichkeitskorrelationen als Korrelation zwischen CDS-Prämien anzunehmen.

Für  $LGD \equiv 1$  ist die Ausfallwahrscheinlichkeit sogar gleich der CDS-Prämie. Zum Beispiel bedeutet eine laufende Prämie von 300 Basispunkten, dass bei einem Kontraktvolumen von 100 Mio. Euro eine Prämienzahlung von 3 Mio. Euro pro Jahr zu leisten ist. Diese Prämie entspricht exakt der Ein-Jahres-Insolvenzwahrscheinlichkeit.

Als Marktportfolio haben wir den Markit-iTraxx Europe verwendet. Dieser Index umfasst die 125 am stärksten gehandelten CDS-Preise in Europa.

Wir werden deshalb die Ausfallwahrscheinlichkeitskorrelation der Banken mittels Panel-Regressionen von CDS-Preisen modellieren und in den folgenden Abschnitten schätzen. Da es in dieser Analyse nicht um Modellbewertung, sondern Modellentwicklung geht, werden verschiedene Modelle hier eingesetzt, um einen plausiblen

Vorschlag zu machen. Uns interessiert vor allem die aus diesen Modellen entstehende Korrelation. Wir werden theoretische Formeln für ihre Berechnungen erstellen und beweisen.

### 3.2 Korrelation von CDS-Preisen

Wie wir im Kapitel 4.1 zeigen werden, liegt ein Zusammenhang zwischen den CDS-Preisen einzelner Banken und den iTraxx-Indizes vor. Dieser Zusammenhang wird mit einer *balancierten* Panel-Regression modelliert, indem alle Variablen reellwertig sind. Ein Vorteil dieses Panel-Regressionstyps ist, dass er eine geschlossene Form von Schätzern durch den *Maximum Likelihood* (ML)-Ansatz erlaubt (vergleiche Demidenko, 2004, S. 49). Die ML-Schätzer können aber auch iterativ durch spezielle Algorithmen berechnet werden. Das Modell ist wie folgt definiert

$$\begin{aligned} CDS_{it} &= \mu + \gamma_i + \eta CDS_{i,t-1} + \beta iTraxx_t + v_{it}, \\ i &= 1, \dots, N, t = 1, \dots, T \end{aligned} \quad (34)$$

und

$$v_{it} = \alpha_i + \lambda_t + \varepsilon_{it}, \quad (35)$$

wobei

- $CDS_{it}$  den  $t$ -ten beobachteten CDS-Preis der Bank  $i$  und  $CDS_{i0}$  den CDS-Preis der Bank  $i$  vor der Beobachtungsphase repräsentiert,
- $iTraxx_t$  den iTraxx-Index aller Finanzmarktunternehmen am Tag  $t$  ist,
- $\mu$  die Mittelwerte der CDS-Preise von Banken im Beobachtungszeitsraum,  $\gamma_i$  der von der Bank  $i$  abhängige feste Effekt und  $\beta$  und  $\eta$  Skalare – die den Einfluss von iTraxx beziehungsweise der verzögerten CDS auf die exogene Variable messen – sind,
- $v_{it}$  (mit Varianz  $\sigma_v^2$ ) den Effekt der ausgelassenen Variable repräsentiert, die zu den Banken und zu dem Zeitraum gehören. Der Fehlerterm  $v_{it}$  besteht aus drei Komponenten:
  - der unbeobachtbare individuenspezifische Effekt  $\alpha_i$ , der aus allen zufälligen zeitinvarianten Effekten der Bank  $i$  besteht,



### 3 Korrelation von Ausfallwahrscheinlichkeiten

- der unbeobachtbare Zeiteffekt  $\lambda_t$ , der aus allen zufälligen individuellen invarianten Effekten besteht
- der Fehlerterm  $\varepsilon_{it}$ , der über Individuen und Zeit variiert.

Es wird zum einen angenommen, dass sowohl  $\lambda_t \sim N(0, \sigma_\lambda^2)$ ,  $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$  und  $\alpha_i \sim N(0, \sigma_\alpha^2)$  unabhängig voneinander sind, als auch dass  $\{iTraxx_t, t = 1 \dots T\}$  nicht stochastisch ist. Zum anderen wird verlangt, dass die Fehlerterme  $v_{it}$  unabhängig voneinander sind und eine Varianz-Kovarianzmatrix  $\Omega$  haben. Die Verzögerte endogene Variable ( $CDS_{i,t-1}$ ), der Marktindex ( $iTraxx_t$ ) sowie  $\gamma_i$  kommen nur alternativ in Modellen vor. In der gesamten Studie ist  $\{CDS_{i0}, i = 1, \dots, N\}$  als fest anzunehmen. Die Modelle (34) und (35) können in Matrix-Vektor-Schreibweise wie folgt umgeschrieben werden

$$CDS = X\delta + v \quad (36)$$

$$v = (I_N \otimes 1_T)\alpha + (1_N \otimes I_T)\lambda + (I_N \otimes I_T)\varepsilon, \quad (37)$$

wobei

- $\otimes$  das Kronecker-Produkt ist,
- $CDS \in \mathbb{R}^{N \cdot T}$  der Vektor der beobachteten CDS-Preise aller Banken ist,
- $X \in \mathbb{R}^{N \cdot T \times m}$  die Designmatrix der unabhängigen Variable ist,
- $\delta \in \mathbb{R}^m$  der Vektor der Bankenparameter,
- $\alpha \in \mathbb{R}^N$  der Vektor der unabhängigen zufälligen Individueneffekte mit Kovarianzmatrix (vergleiche Arnold, 1981, S. 41)

$$\begin{aligned} \Omega(\alpha) &= (I_N \otimes 1_T)\sigma_\alpha^2(I_N' \otimes 1_T') \\ &= \sigma_\alpha^2(I_N \otimes 1_T 1_T') \\ &= \sigma_\alpha^2 \Omega^*(\alpha), \end{aligned}$$

- $\lambda \in \mathbb{R}^T$  der Vektor der unabhängigen zufälligen Zeiteffekte mit Kovarianzmatrix

$$\begin{aligned} \Omega(\lambda) &= (1_N \otimes I_T)\sigma_\lambda^2(1_N' \otimes I_T') \\ &= \sigma_\lambda^2(1_N 1_N' \otimes I_T) \\ &= \sigma_\lambda^2 \Omega^*(\lambda), \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

- $\varepsilon \in \mathbb{R}^{N \cdot T}$  der Vektor der Fehlerterme mit Kovarianzmatrix

$$\begin{aligned}\Omega(\varepsilon) &= (\mathbf{I}_N \otimes \mathbf{I}_T) \sigma_\varepsilon^2 (\mathbf{I}'_N \otimes \mathbf{I}'_T) \\ &= \sigma_\varepsilon^2 (\mathbf{I}_N \otimes \mathbf{I}_T) \\ &= \sigma_\varepsilon^2 \mathbf{I}_{N \cdot T} \\ &= \sigma_\varepsilon^2 \Omega^*(\varepsilon),\end{aligned}$$

- $v$  der gesamte Fehlerterm im Modell mit Kovarianzmatrix

$$\Omega = \Omega(\alpha) + \Omega(\lambda) + \Omega(\varepsilon). \quad (38)$$

Zur Beurteilung der Güte des Modells wird sowohl das *adjustierte Bestimmtheitsmaß* ( $R_{adj}^2$ ) als auch das *Akaike Information Criterion* (AIC) verwendet. Das erste Bestimmtheitsmaß gibt eine Übersicht darüber, wie gut die Modellparameter geschätzt wurden, und hat die folgende Form (vergleiche Draper und Smith, 1998, S. 140)

$$R_{adj}^2 = 1 - \left( 1 - \frac{\hat{\delta}^t X^t CDS - N \cdot T \overline{\overline{CDS}}^2}{CDS^t CDS - N \cdot T \overline{\overline{CDS}}^2} \right) \left( \frac{N \cdot T - 1}{N \cdot T - m} \right),$$

wobei

- $\hat{\delta}$  den geschätzten Parametervektor im Modell repräsentiert,

- $\overline{\overline{CDS}} = \frac{1}{N \cdot T} \sum_{i=1}^N \sum_{t=1}^T CDS_{it}$   
den mittleren CDS-Preis über die Zeit und Banken repräsentiert,

- $CDS = (CDS_{11}, \dots, CDS_{1T}, CDS_{21}, \dots, CDS_{2T}, \dots, CDS_{N1}, \dots, CDS_{N \cdot T})'$   
der Vektor der CDS-Preise aller Banken ist.

Das zweite Bestimmtheitsmaß hängt von der ML-Funktion ab. Für die Modelle (36) und (37) lässt sich die ML-Funktion wie folgt schreiben

$$L(\delta, \Omega) = (2\pi)^{\frac{-N \cdot T}{2}} |\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} (CDS - X\delta)' \Omega^{-1} (CDS - X\delta)\right\}. \quad (39)$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Logarithmiert man diese ML-Funktion, ergibt sich

$$l(\delta, \Omega) = -\frac{1}{2} \left[ N \cdot T \ln(2\pi) + \ln |\Omega| + (CDS - X\delta)' \Omega^{-1} (CDS - X\delta) \right].$$

Maximiert man jetzt die logarithmierte ML-Funktion, indem man sie nach  $\delta$  und  $\Omega$  ableitet und sie gleich Null setzt, ergeben sich die ML-Schätzer  $\hat{\delta}_{ML}$  und  $\hat{\Omega}_{ML}$ . Sei nun

$$l_{max} = l(\hat{\delta}_{ML}, \hat{\Omega}_{ML}),$$

dann hat das *Akaike Information Criterion* die Form (vergleiche Demidenko, 2004, S. 11ff.)

$$AIC = -2l_{max} + 2m.$$

Basierend auf diesem Kriterium können auch verschiedene Modelle anhand der Güte verglichen werden. Generell wird von zwei Modellen dasjenige mit dem niedrigeren AIC ausgewählt.

Um die Korrelation von CDS-Preisen über die Banken und/oder über die Zeit messen zu können, verwenden wir die gewöhnliche Formel

$$\rho_{ij}(t, s) = \text{korr}(CDS_{it}, CDS_{js}) = \frac{\text{kov}(CDS_{it}, CDS_{js})}{\sqrt{\text{var}(CDS_{it})} \sqrt{\text{var}(CDS_{js})}}$$
$$i, j = 1, \dots, N \quad \text{und} \quad t, s = 1, \dots, T$$

Im Folgenden sollen unterschiedliche Annahmen für die Modelle (34) und (35) je nach Ziel getroffen und die resultierenden Modelle untersucht werden.

### 3.2.1 Fester Unternehmenseffekt

Damit keine zufälligen sondern feste Effekte im Modell berücksichtigt werden, setzen wir in diesem Abschnitt

$$\sigma_\alpha^2 = \sigma_\lambda^2 = 0,$$

Dazu sei  $\alpha_i^* = \gamma_i + \mu$  der mittlere CDS-Preis der Bank  $i$  im Beobachtungszeitraum. Wir erhalten das folgende Modell

$$\begin{aligned} CDS_{it} &= \alpha_i^* + \eta CDS_{i,t-1} + \beta iTraxx_t + \varepsilon_{it}, \\ i &= 1, \dots, N, \quad t = 1, \dots, T. \end{aligned} \tag{40}$$

Dieses Modell wird in der englischsprachigen Literatur als *fixed-effects model* geführt. Dabei bezeichnet  $\alpha_i^*$  den *festen Unternehmenseffekt*. Formen wir das Modell in Matrix-Vektor-Schreibweise um, so erhalten wir

$$CDS = X\delta + \varepsilon,$$

wobei

$$\varepsilon = (\varepsilon_{11}, \dots, \varepsilon_{1T}, \varepsilon_{21}, \dots, \varepsilon_{2T}, \dots, \varepsilon_{N1}, \dots, \varepsilon_{N \cdot T})'.$$

Wir werden hier die folgenden Fälle betrachten, die sich durch die Struktur von  $X$  bez.  $\delta$  unterscheiden:

- $\eta = 0$

$$X = \begin{pmatrix} 1_T & 0 & \cdots & 0 & iTraxx \\ 0 & 1_T & \cdots & 0 & iTraxx \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1_T & iTraxx \end{pmatrix} \in \mathbb{R}^{N \cdot T \times (N+1)}$$

$$\delta = (\alpha_1^*, \dots, \alpha_N^*, \beta)' \in \mathbb{R}^{N+1}$$

für das *statische Modell*,

- $\beta \neq 0$  und  $\eta \neq 0$

$$X = \begin{pmatrix} 1_T & 0 & \cdots & 0 & CDS_{1,-1} & iTraxx \\ 0 & 1_T & \cdots & 0 & CDS_{2,-1} & iTraxx \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1_T & CDS_{N,-1} & iTraxx \end{pmatrix} \in \mathbb{R}^{N \cdot T \times (N+2)}.$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

$$\delta = (\alpha_1^*, \dots, \alpha_N^*, \eta, \beta)' \in \mathbb{R}^{N+2}$$

für das *dynamische Modell* mit der Variable *iTraxx*,

- $\beta = 0$  und  $\eta \neq 0$

$$X = \begin{pmatrix} 1_T & 0 & \cdots & 0 & CDS_{1,-1} \\ 0 & 1_T & \cdots & 0 & CDS_{2,-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1_T & CDS_{N,-1} \end{pmatrix} \in \mathbb{R}^{N \cdot T \times (N+1)}.$$

$$\delta = (\alpha_1^*, \dots, \alpha_N^*, \eta)' \in \mathbb{R}^{N+1}$$

für das *dynamische Modell* ohne die Variable *iTraxx*.

Dabei ist

$$iTraxx = (iTraxx_1, \dots, iTraxx_T)',$$

$$CDS_{i,-1} = (CDS_{i,0}, \dots, CDS_{i,T-1})'.$$

Danach wird entschieden, welcher der Fälle vorteilhaft ist. Dies bedeutet, welches der beiden Modelle das niedrigere AIC besitzt. Unter Anwendung der *Kleinsten-Quadrate* (KQ)-Methode erhält man folgende Schätzer für  $\delta$  und  $\sigma_\varepsilon^2$  (vergleiche Chatterjee und Hadi, 1988, S. 40)

$$\hat{\delta}_{KQ} = (X'X)^{-1}X'CDS, \tag{41}$$

$$\hat{\sigma}_{KQ}^2(\varepsilon) = \frac{\sum_{i=1}^N \sum_{t=1}^T (CDS_{it} - \widehat{CDS}_{it})^2}{N \cdot T - m}. \tag{42}$$

Wobei

$$\widehat{CDS}_{it} = \hat{\alpha}_i^* + \hat{\beta}_{KQ} iTraxx_t$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

für das statische Modell und

$$\widehat{CDS}_{it} = \hat{\alpha}_i^* + \hat{\eta}_{KQ} CDS_{i,t-1} + \hat{\beta}_{KQ} iTraxx_t$$

oder

$$\widehat{CDS}_{it} = \hat{\alpha}_i^* + \hat{\eta}_{KQ} CDS_{i,t-1}$$

für die dynamischen Modelle ist, je nachdem ob die exogene Variable  $iTraxx$  in Betracht gezogen wird oder nicht, und die Vorhersage des CDS-Preises der Bank  $i$  zum Zeitpunkt  $t$  repräsentiert.

Unter Annahme des festen Unternehmenseffekts ist das AIC gegeben durch die Form (vergleiche Demidenko, 2004, S. 13)

$$AIC = 2m + N \cdot T \ln \frac{N \cdot T - m}{N \cdot T} \hat{\sigma}_{KQ}^2(\varepsilon).$$

#### Statisches Modell

In diesem Abschnitt wollen wir ausschließlich den Einfluss der  $iTraxx$ -Indizes auf den CDS-Preisen untersuchen. Deshalb wird der Effekt der verzögerten exogenen Variable im Modell (40) isoliert. Dafür setzen wir  $\eta = 0$  und erhalten folgendes Modell

$$\begin{aligned} CDS_{it} &= \alpha_i^* + \beta iTraxx_t + \varepsilon_{it}, \\ i &= 1, \dots, N, \quad t = 1, \dots, T. \end{aligned} \quad (43)$$

Damit erhalten wir explizit folgende Schätzer für  $\alpha_i^*$ ,  $\beta$  (vergleiche Hsiao, 2003, S. 32)

$$\hat{\beta}_{KQ} = \left[ N \sum_{t=1}^T (iTraxx_t - \overline{iTraxx})(iTraxx_t - \overline{iTraxx})' \right]^{-1} \left[ \sum_{i=1}^N \sum_{t=1}^T (iTraxx_t - \overline{iTraxx})(CDS_{it} - \overline{CDS}_i) \right], \quad (44)$$

$$\hat{\alpha}_i^* = \overline{CDS}_i - \hat{\beta}_{KQ} \overline{iTraxx}.$$

Wobei

$$\overline{CDS}_i = \frac{1}{T} \sum_{t=1}^T CDS_{it}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

der mittlere CDS-Preis der Bank  $i$  ist und

$$\overline{iTraxx} = \frac{\sum_{t=1}^T iTraxx_t}{T}$$

der mittlere iTraxx-Index der Finanzmarktunternehmen im Beobachtungszeitsraum ist.

Der Korrelationskoeffizient hat die folgende Form

$$\begin{aligned} \rho_{ij}(t, s) &= \frac{\text{kov}(\varepsilon_{it}, \varepsilon_{js})}{\sqrt{\sigma_\varepsilon^2} \sqrt{\sigma_\varepsilon^2}} \\ &= \begin{cases} 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ sonst.} \end{cases} \end{aligned} \quad (45)$$

Diese theoretischen Ergebnisse der Korrelation würden bedeuten, dass die CDS-Preise von Banken weder über die Zeit noch über Banken korreliert sind. Es bleibt die Frage, ob diese Schlussfolgerung realistisch ist. Die Ergebnisse der empirischen Autokorrelation (Tabelle 2 und Tabelle 3) zeigen, dass alle Banken sowohl über die Zeit als auch unter einander – und zwar mit einer Längsschnittkorrelation von mindestens 0,95 für  $\tau = 1$  sowie einer Querschnittkorrelation von mindestens 0,52 – korreliert sind. Diesem Ergebnis widerspricht das statische Modell mit festem Unternehmenseffekt.

#### Dynamisches Modell

Für das dynamische Modell werden zwei Fälle betrachtet. Zum einem wird das Verhalten der CDS-Preise untersucht, wenn die iTraxx-Indizes und die verzögerten CDS-Preise gleichzeitig im Modell in Betracht gezogen werden ( $\eta \neq 0$ ). Das Modell hat die Form

$$\begin{aligned} CDS_{it} &= \alpha_i^* + \eta CDS_{i,t-1} + \beta iTraxx_t + \varepsilon_{it}, \\ i &= 1, \dots, N, t = 1, \dots, T. \end{aligned} \quad (46)$$

Zum anderen, wird die Variable iTraxx isoliert. Wir konzentrieren uns ausschließlich auf den Einfluss der verzögerten CDS-Preise. Dafür setzen wir  $\beta = 0$  und erhalten folgendes Modell

$$\begin{aligned} CDS_{it} &= \alpha_i^* + \eta CDS_{i,t-1} + \varepsilon_{it}, \\ i &= 1, \dots, N, t = 1, \dots, T. \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Die Modellparameter können auch hier mithilfe der Methode der Kleinsten-Quadrate wie in (41) und (42) ermittelt werden. Im Fall  $\beta = 0$  vereinfachen sich die Parameter  $\hat{\eta}_{KQ}$  und  $\hat{\alpha}_i^*$  zu (vergleiche Hsiao, 2003, S. 71ff.)

$$\begin{aligned}\hat{\eta}_{KQ} &= \frac{\sum_{i=1}^N \sum_{t=1}^T (CDS_{it} - \overline{CDS}_i)(CDS_{i,t-1} - \overline{CDS}_{i,-1})}{\sum_{i=1}^N \sum_{t=1}^T (CDS_{i,t-1} - \overline{CDS}_{i,-1})^2} \\ &= \eta + \frac{\frac{1}{N \cdot T} \sum_{i=1}^N \sum_{t=1}^T (CDS_{i,t-1} - \overline{CDS}_{i,-1})(\varepsilon_{i,t} - \bar{\varepsilon}_i)}{\frac{1}{N \cdot T} \sum_{i=1}^N \sum_{t=1}^T (CDS_{i,t-1} - \overline{CDS}_{i,-1})^2},\end{aligned}$$

$$\hat{\alpha}_i^* = \overline{CDS}_i - \hat{\eta}_{KQ} \overline{CDS}_{i,-1},$$

wobei

$$\overline{CDS}_{i,-1} = \frac{1}{T} \sum_{t=1}^T CDS_{i,t-1} \quad \text{und} \quad \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}.$$

Diese geschätzten Parameter sind inkonsistent, denn die asymptotische Verzerrung ist in diesem Modell wie folgt gegeben (vergleiche Hsiao, 2003, S. 73ff.) und (Arellano, 2003, S. 84f.)

$$\text{plim}_{N \rightarrow \infty}(\hat{\eta}_{KQ} - \eta) = -\frac{\frac{1+\eta}{T-1} \left[ 1 - \frac{1}{T} \frac{1-\eta^T}{1-\eta} \right]}{1 - \frac{2\eta}{(1-\eta)(T-1)} \left[ 1 - \frac{1-\eta^T}{T(1-\eta)} \right]}.$$

Aus dieser Formel ergibt sich, dass der Schätzer  $\hat{\eta}_{KQ}$  nur für  $T \rightarrow \infty$  konsistent ist. Für niedrige Werte von  $T$  ist diese Verzerrung bedeutend. Zum Beispiel haben wir für den Fall  $T = 2$  die folgende Verzerrung

$$\text{plim}_{N \rightarrow \infty}(\hat{\eta}_{KQ} - \eta) = -\frac{1+\eta}{2}.$$

Dies würde bedeuten, dass die Schätzmethode beim dynamischen Modell mit Unternehmenseffekten im Allgemeinen nicht effizient ist, wenn  $T$  nicht genug groß ist, da die geschätzten Parameter dort inkonsistent sind.

**Lemma 3.1.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur im Modell (46) durch die folgende Korrelationsfunktion ableiten.

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \hat{\eta}^{|t-s|} \sqrt{\frac{1-\hat{\eta}^{2\min(t,s)}}{1-\hat{\eta}^{2\max(t,s)}}} & , \quad \text{für } i = j, \\ 0 & , \quad \text{sonst.} \end{cases} \quad (47)$$



### 3 Korrelation von Ausfallwahrscheinlichkeiten

**Beweis.** Schreibt man das Modell (46) wie folgt um (vergleiche Arellano, 2003, S. 83),

$$CDS_{it} = \eta^t CDS_{i,0} + \sum_{r=1}^t \eta^{t-r} (\alpha_i^* + \beta i Traxx_r + \varepsilon_{ir}),$$

$$i = 1, \dots, N, t = 1, \dots, T$$

so erhält man

$$\begin{aligned} \text{var}(CDS_{it}) &= \sum_{r=1}^t \text{var} [\eta^{t-r} (\alpha_i^* + \beta i Traxx_r + \varepsilon_{ir})], \\ &= \sum_{r=1}^t \eta^{2(t-r)} \sigma_\varepsilon^2 \\ &= \sigma_\varepsilon^2 \frac{1 - \eta^{2t}}{1 - \eta^2}. \end{aligned}$$

Sei  $s, t \in \{1, 2, \dots, T\}$ . Wir nehmen ohne Beschränkung der Allgemeinheit an, dass gilt  $t \geq s$ . Dann folgt

$$\begin{aligned} \text{kov}(CDS_{it}, CDS_{js}) &= \text{kov} \left( \sum_{r=1}^t \eta^{2(t-r)} \varepsilon_{ir}, \sum_{q=1}^s \eta^{2(s-r)} \varepsilon_{jq} \right) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \text{kov}(\varepsilon_{ir}, \varepsilon_{jq}) \\ &= \begin{cases} \sigma_\varepsilon^2 \eta^{t-s} \sum_{q=1}^s \eta^{2(s-q)} & , \text{ für } i = j, \\ 0 & , \text{ sonst,} \end{cases} \\ &= \begin{cases} \sigma_\varepsilon^2 \eta^{t-s} \frac{1 - \eta^{2s}}{1 - \eta^2} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases} \end{aligned} \quad (48)$$

□

Die Formel (47) ist zwar eine Verbesserung von (45), da die Korrelationsfunktion in diesem Fall mit wachsendem Abstand zwischen  $t$  und  $s$  sinkt, muss aber noch verbessert werden, da die Querschnittkorrelation immer Null geblieben ist.

Im weiteren Verlauf werden folgende Annahmen getroffen

$$\gamma_i = 0, i = 1, \dots, N. \quad (49)$$

### 3.2.2 Zufälliger Unternehmenseffekt

Zur Verbesserung des Modells mit festen Unternehmenseffekten wird in (34) und (35) zusätzlich zu der Gleichung (49) angenommen, dass gilt  $\sigma_\lambda^2 = 0$ . Der Anteil  $\alpha_i \sim N(0, \sigma_\alpha^2)$  wird als *zufälliger Unternehmenseffekt* bezeichnet. Der unternehmensspezifische Effekt  $\alpha_i$  ist auch dafür zuständig, dass die CDS-Preise eines Unternehmens über die Zeit korreliert sind. Dies ist plausibel, wenn man bedenkt, dass im Allgemeinen dasselbe Management über einen längeren Zeitraum zu dem unternehmerischen Erfolg beiträgt.

Die Modellierung der Fehlerterme sieht wie folgt aus:

$$v_{it} = \alpha_i + \varepsilon_{it}.$$

Und damit hat das gesamte Modell die folgende Form

$$\begin{aligned} CDS_{it} &= \mu + \eta CDS_{i,t-1} + \beta iTraxx_t + \alpha_i + \varepsilon_{it}, \\ i &= 1 \dots N, t = 1, \dots, T. \end{aligned} \quad (50)$$

Wobei

$$\begin{aligned} E(\alpha_i) &= E(\varepsilon_{it}) = E(\alpha_i iTraxx_t) = E(\varepsilon_{it} iTraxx_t) \\ &= E(\alpha_i CDS_{i0}) = E(\varepsilon_{it} CDS_{i,t-1}) = 0, \end{aligned} \quad (51)$$

$$E(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2 & , \text{ für } i = j, \\ 0 & , \text{ sonst,} \end{cases} \quad (52)$$

$$E(\varepsilon_{it} \varepsilon_{js}) = \begin{cases} \sigma_\varepsilon^2 & , \text{ für } i = j, t = s, \\ 0 & , \text{ sonst.} \end{cases} \quad (53)$$

Formt man das Modell (50) in Matrix-Vektor-Schreibweise um, so erhält man

$$\begin{aligned} CDS_i &= X_i \delta + 1_T \alpha_i + \varepsilon_i, \\ i &= 1, \dots, N. \end{aligned} \quad (54)$$

Auch hier sollen die Einflüsse der verzögerten endogenen Variable und die Variable  $iTraxx$  auf die CDS-Preise im Modell (50) analysiert werden. Dafür isolieren wir aufeinanderfolgend die Effekte dieser Variablen im Modell. Dabei setzen wir

### 3 Korrelation von Ausfallwahrscheinlichkeiten

- $\eta = 0$

$$\begin{aligned} X_i &= (1_T \quad iTraxx) \in \mathbb{R}^{T \times 2}, \\ i &= 1, \dots, N, \\ \delta &= (\mu, \beta) \in \mathbb{R}^2, \end{aligned}$$

für das *statische Modell mit zufälligem Unternehmenseffekt*. Wir werden beobachten, wie die Variable *iTraxx* ausschließlich die Ergebnisse der Modellierung mit zufälligem Unternehmenseffekt beeinflusst.

- $\eta \neq 0$

für das *dynamische Modell mit zufälligem Unternehmenseffekt*. Hier ist zu beobachten, wie groß der Einfluss der verzögerten CDS-Preise auf die Modellierung ist. Wir untersuchen dabei sowohl den Einfluss kombinierter Variablen im Modell, indem wir

- $\beta \neq 0$

$$\begin{aligned} X_i &= (1_T \quad CDS_{i,-1} \quad iTraxx) \in \mathbb{R}^{T \times 3}, \\ i &= 1, \dots, N \\ \delta &= (\mu, \eta, \beta) \in \mathbb{R}^3, \end{aligned}$$

setzen als auch ausschließlich die Auswirkung der verzögerten CDS-Preise auf die Ergebnisse der Modellierung, indem wir

- $\beta = 0$

$$\begin{aligned} X_i &= (1_T \quad CDS_{i,-1}) \in \mathbb{R}^{T \times 2}, \\ i &= 1, \dots, N, \\ \delta &= (\mu, \eta) \in \mathbb{R}^2, \end{aligned}$$

setzen. Die Bestimmung der Modellparameter erfolgt in diesem Abschnitt mithilfe des ML-Ansatzes. Ausnutzung von (39) führt zu der folgenden ML-Funktion (vergleiche Hsiao, 2003, S. 78ff.)

$$\begin{aligned} L(\delta, \sigma_\varepsilon^2, \sigma_\alpha^2) &= (2\pi)^{-\frac{N \cdot T}{2}} |V_i|^{-\frac{N}{2}} \times \\ &\exp \left[ -\frac{1}{2} \sum_{i=1}^N (CDS_i - X_i \delta)' V_i^{-1} (CDS_i - X_i \delta) \right]. \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Wobei

$$\begin{aligned}
 V_i &= \sigma_\varepsilon^2 \mathbf{I}_T + \sigma_\alpha^2 \mathbf{1}_T \mathbf{1}_T', \\
 &= \sigma_\varepsilon^2 [\mathbf{I}_T + d \mathbf{1}_T \mathbf{1}_T'], \\
 &= \sigma_\varepsilon^2 V \\
 &\quad \text{mit } i = 1, \dots, N,
 \end{aligned} \tag{55}$$

und

$$d = \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2}. \tag{56}$$

Die Modellparameter können mittels der *Restringierten Maximum Likelihood* (REML)-Methode (vergleiche Demidenko, 2004, S. 58ff.) geschätzt werden, da die gewöhnliche ML-Methode die Varianzkomponenten ( $\sigma_\varepsilon^2$ ,  $\sigma_\alpha^2$ ) unterschätzt, sodass sie verzerrt geschätzt werden (Kiefer, 1980). Die entsprechende REML-Funktion hat die folgende Form (vergleiche Demidenko, 2004, S. 60)

$$\begin{aligned}
 l_R(\delta, \sigma_\varepsilon^2, d) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\varepsilon^2 + \ln \left| \sum_{i=1}^N X_i' V^{-1} X_i \right| \right] + \\
 &\quad -\frac{1}{2} \sum_{i=1}^N \ln |V| + \frac{-1}{2\sigma_\varepsilon^2} \sum_{i=1}^N (CDS_i - X_i \delta)' V^{-1} (CDS_i - X_i \delta).
 \end{aligned} \tag{57}$$

Die REML-Schätzer für  $\delta$ ,  $\sigma_\varepsilon^2$  und  $\sigma_\alpha^2$  können berechnet werden mittels

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N \cdot T - m} \sum_{i=1}^N (CDS_i - X_i^* \delta)' V^{-1} (CDS_i - X_i^* \delta) \tag{58}$$

und Maximierung der resultierenden Funktion, genannt *restringierter Profile-log-likelihood* ( $l_{R,p}$ ), über  $\delta$  und  $d$

$$\begin{aligned}
 l_{R,p}(\delta, d) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \left( \sum_{i=1}^N (CDS_i - X_i \delta)' V^{-1} (CDS_i - X_i \delta) \right) \right] \\
 &\quad -\frac{1}{2} \left[ \ln \left| \sum_{i=1}^N X_i' V^{-1} X_i \right| + \sum_{i=1}^N \ln |V| \right] + const,
 \end{aligned} \tag{59}$$

wobei  $const = (N \cdot T - m) [\ln(N \cdot T - m) - 1]$ .

Für den Fall  $X_i = X_j$ ,  $i, j = 1, \dots, N$  werden wir das Optimierungsproblem analytisch im nächsten Abschnitt lösen. Für den anderen ( $X_i \neq X_j$ ,  $i \neq j$ ) lässt sich das Problem mittels Maximierungsalgorithmen wie Newton-Raphson (NR), Expectation-Maximization (EM), Fisher Scoring (FS) iterativ bestimmen (vergleiche beispielweise Demidenko, 2004, S. 80ff.).

### 3 Korrelation von Ausfallwahrscheinlichkeiten

#### Statisches Modell

Das Ziel ist hier, ausschließlich den Einfluss der Variable *iTraxx* auf die Zielvariable im Modell mit zufälligem Unternehmenseffekt zu ermitteln. Deshalb wird der Effekt der verzögerten CDS-Preise ( $\eta = 0$ ) im Modell isoliert. Das Modell sieht folgendermaßen aus

$$\begin{aligned} CDS_{it} &= \mu + \beta iTraxx_t + \alpha_i + \varepsilon_{it}, \\ i &= 1, \dots, N, \quad t = 1, \dots, T. \end{aligned} \quad (60)$$

Wobei  $\alpha_i$  eine Zufallsvariable ist, die die Abweichung zwischen dem mittleren CDS-Preis ( $\mu$ ) der gesamten Bank und dem von der Bank  $i$  ist.

Unter dieser Modellspezifikation haben wir

$$X_i = (1_T, iTraxx) = X^*.$$

Leitet man  $l_{R,p}(\beta, d)$  nach  $\beta$  und  $d$  ab und setzt die Ableitungen gleich Null, ergibt sich nach einigen Umformulierungen (vergleiche Demidenko, 2004, S. 71f.):

$$\hat{\delta}_{REML} = \left[ \sum_{i=1}^N X_i' V_i^{-1} X_i \right]^{-1} \left[ \sum_{i=1}^N X_i' V_i^{-1} CDS_i \right], \quad (61)$$

$$\hat{d}_{REML} = \frac{1}{(N-1)\hat{\sigma}_{ML}^2(\varepsilon)} \sum_{i=1}^N (\overline{CDS}_i - \overline{\overline{CDS}})^2 - \frac{1}{T}. \quad (62)$$

Wobei

$$\overline{\overline{CDS}} = \frac{1}{N \cdot T} \sum_{i=1}^N \sum_{t=1}^T CDS_{it},$$

und

$$\hat{\sigma}_{ML}^2(\varepsilon) = \frac{1}{N(T-1)} \sum_{i=1}^N \left[ (CDS_i - X^* \delta)' (CDS_i - X^* \delta) - T(\overline{CDS}_i - \overline{\overline{CDS}})^2 \right].$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Es gilt weiter (vergleiche Hsiao, 2003, S. 34–r38)

$$\begin{aligned}
 \hat{\beta}_{REML} &= \left[ \frac{N}{T} iTraxx' QiTraxx + \hat{\psi} N (\overline{iTraxx} - \overline{\overline{iTraxx}})^2 \right]^{-1} \\
 &\quad \left[ \frac{1}{T} \sum_{i=1}^N iTraxx' QCDS_i + \hat{\psi} \sum_{i=1}^N (\overline{iTraxx} - \overline{\overline{iTraxx}}) (\overline{CDS}_i - \overline{\overline{CDS}}) \right] \\
 &= \left[ \frac{N}{T} iTraxx' QiTraxx \right]^{-1} \left[ \frac{1}{T} \sum_{i=1}^N iTraxx' QCDS_i \right] \\
 &= \hat{\beta}_{KQ}, \tag{63}
 \end{aligned}$$

$$\hat{\mu}_{REML} = \overline{\overline{CDS}} - \hat{\beta}_{REML} \overline{iTraxx} = \overline{\overline{CDS}} - \hat{\beta}_{KQ} \overline{iTraxx} = \hat{\mu}_{KQ}, \tag{64}$$

wobei

$$Q = I_T - \frac{1}{T} \mathbf{1}_T \mathbf{1}'_T$$

die Matrix repräsentiert, die den Abstand zwischen einem T-dimensionalen Vektor und seinem Mittelwert angibt,

$$\psi = \frac{\sigma^2(\varepsilon)}{\sigma^2(\varepsilon) + T\sigma^2(\alpha)},$$

$$\overline{\overline{iTraxx}} = \frac{1}{N \cdot T} \sum_{i=1}^N \sum_{t=1}^T iTraxx_t = \overline{iTraxx}.$$

Die durch die Maximierung gefundenen Lösungen  $\hat{\delta}_{REML}$  und  $\hat{d}_{REML}$  setzt man in  $\hat{\sigma}_\varepsilon^2$  ein (vgl. (58)).

Somit erhält man

$$\hat{\sigma}_{REML}^2(\varepsilon) = \sum_{i=1}^N \frac{\|\hat{\varepsilon}_i\|^2 - T(\overline{CDS}_i - \overline{\overline{CDS}})^2}{N(T-1) - m + 1}, \tag{65}$$

wobei

$$\|\hat{\varepsilon}_i\|^2 = (CDS_i - X^* \hat{\delta}_{REML})' (CDS_i - X^* \hat{\delta}_{REML}),$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

die Norm der REML-Residuen repräsentiert.

Die Schätzer (63) und (64) machen klar, dass sich der REML-Schätzer für  $\delta$  unabhängig von der Fehlerkovarianzmatrix bestimmen lässt. Es folgt daraus, dass die Anwendung der KQ-Methode auf das Modell (60) zum gleichen Schätzer für den Parametervektor ( $\delta$ ) wie die REML-Methode führt.

Werden nun die Ausdrücke für  $\hat{\sigma}_{REML}^2(\varepsilon)$  und  $\hat{d}_{REML}$  in die Formel 56 eingesetzt, so erhält man

$$\hat{\sigma}_{REML}^2(\alpha) = \frac{1}{(N-1)} \sum_{i=1}^N (\overline{CDS}_i - \overline{\overline{CDS}})^2 - \frac{\hat{\sigma}_{REML}^2(\varepsilon)}{T}. \quad (66)$$

Als Maße für die Abhängigkeit definieren wir die folgende Korrelationsfunktion

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_{\alpha}^2}{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_{\varepsilon}^2} & , \text{ für } i = j, t \neq s, \\ 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ für } i \neq j. \end{cases} \quad (67)$$

Es wird ersichtlich, dass die Korrelation über die Zeit aufgrund der Annahmen an die Fehlerstruktur konstant bleibt. Dieses Ergebnis scheint aber unrealistisch zu sein, da die empirischen Autokorrelationen zeigen, dass diese Werte mit steigenden Abständen zwischen den Handelstagen sinken. Außerdem beträgt die Korrelation über Banken Null, was die empirische Querschnittkorrelation (Tabelle 3) nicht bestätigt. Im nächsten Abschnitt werden wir untersuchen, wie die CDS-Preise und somit die theoretischen Korrelationen auf das Einfügen der verzögerten endogenen Variable im statischen Modell mit zufälligen Unternehmenseffekten reagieren.

#### Dynamisches Modell

Hier untersuchen wir den Einfluss der verzögerten CDS-Preise bei der Modellierung mit zufälligem Unternehmenseffekt. Dabei stellen wir fest, dass gilt  $\eta \neq 0$ . Wir definieren dazu folgende Modelle, je nachdem, ob wir gleichzeitig den Einfluss von iTraxx ermitteln wollen:

$$\begin{aligned} CDS_{it} &= \mu + \eta CDS_{i,t-1} + \beta iTraxx_t + \alpha_i + \varepsilon_{it}, \\ i &= 1, \dots, N, t = 1, \dots, T \end{aligned} \quad (68)$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

oder ob wir ausschließlich den Einfluss der verzögerten exogenen Variable berücksichtigen möchten:

$$\begin{aligned} CDS_{it} &= \mu + \eta CDS_{i,t-1} + \alpha_i + \varepsilon_{it}, \\ i &= 1, \dots, N, \quad t = 1, \dots, T. \end{aligned} \quad (69)$$

Die Modellparameter für beide Spezifikationen –  $\beta \neq 0$  oder  $\beta = 0$  – lassen sich durch die FS-, NR-, oder EM-Algorithmen schätzen. Aufgrund der vorliegenden Korrelation zwischen  $\alpha_i$  und  $CDS_{i,t-1}$  ist zu erwarten, dass diese Schätzer inkonsistent sind. Aber bei einer umfangreichen Stichprobe, also  $T \rightarrow \infty$  oder  $N \rightarrow \infty$  können die Schätzer effizient sein (vergleiche Hsiao, 2003, S. 78ff.).

**Lemma 3.2.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in Modellen (68) und (69) durch die folgende Korrelationsfunktion ableiten.

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\frac{(1-\eta^t)(1-\eta^s)}{(1-\eta)^2} + \frac{\hat{\sigma}_\varepsilon^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2}}{\hat{\sigma}_\alpha^2}}{\sqrt{\left(\frac{(1-\eta^s)}{1-\eta}\right)^2 + \frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_\alpha^2} \frac{1-\eta^{2s}}{1-\eta^2}} \sqrt{\left(\frac{(1-\eta^t)}{1-\eta}\right)^2 + \frac{\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_\alpha^2} \frac{1-\eta^{2t}}{1-\eta^2}}}, & \text{für } i = j, \\ 0 & \text{sonst,} \end{cases} \quad (70)$$

*Beweis.* Aufgrund der Annahmen an  $CDS_{i,0}$  transformieren wir das Modell (68) wie folgt, um die Varianz-Kovarianzfunktion der CDS-Preise leicht zu berechnen.

$$\begin{aligned} CDS_{it} &= \eta^t CDS_{i,0} + \sum_{r=1}^t \eta^{t-r} (\mu + \alpha_i + \beta i \text{Traxx}_r + \varepsilon_{ir}), \\ i &= 1, \dots, N, \quad t = 1, \dots, T \end{aligned} \quad (71)$$

Sei  $s, t \in \{1, 2, \dots, T\}$ . Wir nehmen ohne Beschränkung der Allgemeinheit an, dass gilt  $t \geq s$ . Dann folgt

$$\begin{aligned} Kov_{ij}(t, s) &:= kov(CDS_{it}, CDS_{js}) \\ &= kov\left(\sum_{r=1}^t \eta^{t-r} (\alpha_i + \varepsilon_{ir}), \sum_{q=1}^s \eta^{s-q} (\alpha_j + \varepsilon_{jq})\right) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\alpha_i + \varepsilon_{ir}, \alpha_j + \varepsilon_{jq}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} [kov(\alpha_i, \alpha_j) + kov(\varepsilon_{ir}, \varepsilon_{jq})]. \end{aligned} \quad (72)$$



### 3 Korrelation von Ausfallwahrscheinlichkeiten

Wegen (51), (52) und (53) folgt

$$\begin{aligned}
 \text{Kov}_{ij}(t, s) &= \begin{cases} \sigma_{\alpha}^2 \left( \sum_{r=1}^t \sum_{q=1}^s \eta^{(t-r)+(s-q)} \right) + \sigma_{\varepsilon}^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} & , \text{ für } i = j, \\ 0 & , \text{ sonst,} \end{cases} \\
 &= \begin{cases} \sigma_{\alpha}^2 \frac{1-\eta^t}{1-\eta} \frac{1-\eta^s}{1-\eta} + \sigma_{\varepsilon}^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases} \quad (73)
 \end{aligned}$$

Daraus ergibt sich folgendes für die die Varianz des CDS-Preises einer Bank am Tag  $t$

$$\text{var}(CDS_{js}) = \sigma_{\alpha}^2 \left( \frac{1-\eta^s}{1-\eta} \right)^2 + \sigma_{\varepsilon}^2 \frac{1-\eta^{2s}}{1-\eta^2}.$$

□

Es ist ersichtlich, dass die verzögerten exogenen Variablen die Korrelation beeinflussen. Trotz dieses Einflusses bleibt die Querschnittskorrelation Null. Dies liegt weiter in den Annahmen an die Fehler.

Der ML-Ansatz fordert, dass die Fehlerterme und die zufälligen Effekte normalverteilt sind. Deswegen ist dieser Ansatz nicht robust gegen Verteilungsannahmen. Die quadratische Schätzung der Varianzparameter scheint dagegen adäquat im linearen Modell. Zum einen, weil das Verfahren keine iterative Lösung erfordert, zum anderen, weil die daraus resultierenden Schätzer ohne Annahmen über die Verteilung unverzerrt sind. Dies ist gerade in unserem Fall interessant, da die Annahmen über die Normalverteilung häufig verletzt sind, sodass die ML-Schätzung der Kovarianzmatrix  $\Omega$  nicht mehr effizient ist. Die Varianzschätzung ist im Allgemeinen entscheidend für den Parameter  $\delta$ , da bei bekannten  $\Omega$   $\delta$  mithilfe der verallgemeinerten KQ-Methode ermittelt werden kann. Zu der Klasse der quadratischen unverzerrten Schätzern gehören *Minimum Norm Quadratic Unbiased* (MINQU)-Schätzer für die Varianzparameter, die von Rao eingeführt wurden. Für die Modelle mit Annahmen über die Normalverteilung führt dieser Ansatz zu dem quadratischen unverzerrten Schätzer mit minimaler Varianz (vergleiche Demidenko, 2004, S. 157–158). Demidenko, 2004, S. 177 zeigt, dass die REML-Schätzer von Varianzkomponenten  $(\sigma_{\varepsilon}^2, \sigma_{\alpha}^2)$  im balancierten Modell mit zufälligem Unternehmenseffekt gleich den MINQU-Schätzern sind.

### AR(1)-Fehler

Wir wollen untersuchen, ob die Reststörgrößen ( $\varepsilon_{it}$ ) die Struktur eines AR(1)-Prozesses im Modell mit zufälligem Unternehmenseffekt – das heißt, die Modelle (60), (68) sowie (69) – aufweisen. Wir modellieren die Reststörgrößen  $\varepsilon_{it}$  wie folgt

$$\varepsilon_{it} = \phi \varepsilon_{i,t-1} + \zeta_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \quad |\phi| < 1, \quad (74)$$

wobei sowohl  $\zeta_{it} \sim N(0, \sigma_\zeta^2)$  unabhängig identisch verteilt als auch unabhängig von  $\varepsilon_{i,t-1}$  sind. Wir erhalten somit die folgende Struktur für die Störgrößen.

$$\begin{aligned} v_{it} &= \alpha_i + \varepsilon_{it}, \\ i &= 1, \dots, N, \quad t = 1, \dots, T, \end{aligned}$$

und

$$\begin{aligned} \varepsilon_{it} &= \phi \varepsilon_{i,t-1} + \zeta_{it}, \\ |\phi| &< 1. \end{aligned}$$

Die Kovarianzfunktion für  $\varepsilon_{it}$  hat dann die Form (vergleiche Fuller, 1976, S. 37)

$$\text{kov}(\varepsilon_{it}, \varepsilon_{is}) = \sigma_\zeta^2 \frac{\phi^{|t-s|}}{1 - \phi^2}. \quad (75)$$

Daraus ergibt sich – mithilfe von (4) – die folgende Varianzkovarianzmatrix für jedes Unternehmen.

$$V_i^* = \sigma_\zeta^2 (V(\phi) + d^* \mathbf{1}_T \mathbf{1}_T'), \quad i = 1, \dots, N, \quad (76)$$

wobei

$$d^* = \frac{\sigma_\alpha^2}{\sigma_\zeta^2}. \quad (77)$$

Mithilfe dieser Kovarianzmatrix erhalten wir die folgende REML-Funktion (vergleiche Demidenko, 2004, S. 60ff.):

$$\begin{aligned} l_R^*(\delta, \sigma_\zeta^2, \phi, d^*) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\zeta^2 + \ln \left| \sum_{i=1}^N X_i' V^{*-1} X_i \right| + N \ln |V^*| \right] \\ &\quad - \frac{1}{2\sigma_\zeta^2} \sum_{i=1}^N (CDS_i - X_i \delta)' V^{*-1} (CDS_i - X_i \delta). \end{aligned} \quad (78)$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Unter Verwendung von

$$\hat{\sigma}_\zeta^2 = \frac{1}{N \cdot T - m} \sum_{i=1}^N (CDS_i - X_i \delta)' V^{*-1} (CDS_i - X_i \delta) \quad (79)$$

in der REML-Funktion (78), erhält man die folgende restringierte Profile-log-Likelihood-Funktion

$$l_{R,p}^*(\delta, \phi, d^*) = -\frac{1}{2} \left[ (N \cdot T - 2) \ln \left( \sum_{i=1}^N (CDS_i - X_i \delta)' V^{*-1} (CDS_i - X_i \delta) \right) \right] - \frac{1}{2} \left[ \ln \left| \sum_{i=1}^N X_i' V^{*-1} X_i \right| + N \ln |V^*| \right] + const. \quad (80)$$

Es gilt weiter (vergleiche Demidenko, 2004, S. 227ff.):

$$\begin{aligned} Korr^{-1} &= \frac{1}{1 - \phi^2} \begin{pmatrix} 1 & -\phi & 0 & \cdots & \cdots & \cdots & 0 \\ -\phi & 1 + \phi^2 & -\phi & \ddots & & & \vdots \\ 0 & -\phi & 1 + \phi^2 & -\phi & \ddots & & \vdots \\ \vdots & \ddots & -\phi & 1 + \phi^2 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & -\phi & 0 \\ \vdots & & & \ddots & -\phi & 1 + \phi^2 & -\phi \\ 0 & \cdots & \cdots & \cdots & 0 & -\phi & 1 \end{pmatrix} \\ &= \frac{1}{1 - \phi^2} (I_T + \phi^2 I^* - \phi H) \\ &= \frac{1}{1 - \phi^2} U \end{aligned}$$

$$|Korr| = 1 - \phi^2,$$

$$|V^*| = \frac{1 - \phi^2}{\tilde{d}} \left[ \tilde{d} + T + \phi^2(T - 2) - 2\phi(T - 1) \right],$$

$$V^{*-1} = \frac{1}{1 - \phi^2} \left[ U - U 1_T (\tilde{d} + 1_T' U 1_T)^{-1} 1_T' U \right]$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

wobei  $H$  und  $I^*$   $T \times T$  Matrizen sind, die wie folgt definiert werden:

$$H = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \vdots \\ 0 & 1 & 0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}, \quad I^* = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 0 \end{pmatrix}$$

und

$$\tilde{d} = (1 - \phi^2) / d^*. \quad (81)$$

Werden die Ausdrücke für  $|V^*|$ ,  $V^{*-1}$  in die obige restringierte Profile-Likelihood eingesetzt und einige Umformulierungen durchgeführt, reduziert sich die restringierte Profile-Likelihood zu

$$l_{R,p}^*(\delta, \phi, \tilde{d}) = -\frac{(N \cdot T - 2)}{2} \ln \left( s_0 + \phi^2 s_1 - \phi s_2 - \sum_{i=1}^N \frac{r_i^2}{g} \right) - \frac{1}{2} \left[ N \ln g - (N \cdot T - 2 - N) \ln(1 - \phi^2) - N \ln \tilde{d} + \ln f \right] + const, \quad (82)$$

wobei

$$g = \tilde{d} + T + \phi^2(T - 2) - 2\phi(T - 1),$$

$$f = \left( \frac{N}{1 - \phi^2} \right)^2 \det \begin{pmatrix} 1'_T U 1_T - \frac{(1'_T U 1_T)^2}{g} & 1'_T U i Traxx - \frac{(1'_T U 1_T 1'_T U i Traxx)}{g} \\ i Traxx' U 1_T - \frac{i Traxx' U 1_T 1'_T U 1_T}{g} & i Traxx' U i Traxx - \frac{(i Traxx' U 1_T)^2}{g} \end{pmatrix},$$

$$r_i = \sum_{t=1}^T e_{i,t} + \phi^2 \sum_{t=2}^{T-1} e_{i,t} - \phi \left( 2 \sum_{t=2}^{T-1} e_{i,t} + e_{i,1} + e_{i,T} \right),$$

$$s_0 = \sum_{t=1}^T e_{i,t}^2, s_1 = \sum_{t=2}^{T-1} e_{i,t}^2, s_2 = 2 \sum_{t=1}^{T-1} e_{i,t} e_{i,t+1}.$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Mithilfe der oben genannten Maximierungsalgorithmen lässt sich die restringierte Profile-Likelihood-Funktion ( $l_{R,p}^*$ ) über  $\delta$ ,  $\phi$ ,  $\vec{d}$  maximieren. Setzt man die durch das Maximierungsproblem gefundenen Schätzer  $\hat{\delta}_{REML}^*$ ,  $\hat{\phi}$ ,  $\hat{\vec{d}}$  in (81) ein, so erhält man den Schätzer  $\hat{d}^*$ , der selbst in (76) eingesetzt wird, sodass der Schätzer  $\hat{V}^*$  ermittelt werden kann. Mithilfe von (79) lässt sich die Fehlervarianz  $\sigma_{REML}^2(\zeta)$  und daraus  $\sigma_{REML}^2(\alpha)$  schätzen.

**Lemma 3.3.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur im statischen Modell mit zufälligem Unternehmenseffekt und vorliegendem AR(1)-Prozess in Reststögrößen durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{(1-\hat{\phi}^2)\hat{\sigma}_\alpha^2 + \hat{\phi}^{|t-s|}\hat{\sigma}_\zeta^2}{(1-\hat{\phi}^2)\hat{\sigma}_\alpha^2 + \hat{\sigma}_\zeta^2} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases}$$

*Beweis.* Wegen (75) hat die Kovarianzfunktion des gesamten Fehlerterms die Form

$$\text{kov}(v_{it}, v_{js}) = \begin{cases} \sigma_\alpha^2 + \sigma_\zeta^2 \frac{\phi^{|t-s|}}{1-\phi^2} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases}$$

Es folgt daraus, dass gilt

$$\text{kov}(v_{it}, v_{js}) = \begin{cases} \frac{\sigma_\alpha^2 + \sigma_\zeta^2 \frac{\phi^{|t-s|}}{1-\phi^2}}{\sigma_\alpha^2 + \frac{\sigma_\zeta^2}{1-\phi^2}} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases}$$

□

**Lemma 3.4.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in dynamischen Modellen mit zufälligem Unternehmenseffekt und vorliegendem AR(1)-Prozess in Reststögrößen durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_\alpha^2 \frac{1-\hat{\eta}^t}{1-\hat{\eta}} \frac{1-\hat{\eta}^s}{1-\hat{\eta}} + \frac{\hat{\sigma}_\zeta^2}{1-\hat{\phi}^2} \sum_{r=1}^t \sum_{q=1}^s \hat{\eta}^{t-r+s-q} \hat{\phi}^{|r-q|}}{\sqrt{\hat{v}\text{ar}(CDS_{it})} \sqrt{\hat{v}\text{ar}(CDS_{is})}} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

wobei

$$\hat{v}ar(CDS_{it}) = \hat{\sigma}_\alpha^2 \left( \frac{1 - \hat{\eta}^t}{1 - \hat{\eta}} \right)^2 + \frac{\hat{\sigma}_\zeta^2}{1 - \hat{\phi}^2} \sum_{r=1}^t \sum_{q=1}^t \hat{\eta}^{2t-r-q} \hat{\phi}^{|r-q|}.$$

**Beweis.** Wegen

$$CDS_{it} = \eta^t CDS_{i,0} + \sum_{r=1}^t \eta^{t-r} (\mu + \beta i Traxx_r + v_{ir}),$$

lässt sich die Kovarianzfunktion wie folgt berechnen

$$\begin{aligned} Kov_{ij}(t, s) &= kov(CDS_{it}, CDS_{js}) \\ &= kov \left( \sum_{r=1}^t \eta^{t-r} v_{ir}, \sum_{q=1}^s \eta^{s-q} v_{jq} \right) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(v_{ir}, v_{jq}) \\ &= \begin{cases} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \left[ \sigma_\alpha^2 + \sigma_\zeta^2 \frac{\phi^{|r-q|}}{1-\phi^2} \right] & , \text{ für } i = j, \\ 0 & , \text{ sonst,} \end{cases} \\ &= \begin{cases} \sigma_\alpha^2 \frac{1-\eta^t}{1-\eta} \frac{1-\eta^s}{1-\eta} + \frac{\sigma_\zeta^2}{1-\phi^2} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \phi^{|r-q|} & , \text{ für } i = j, \\ 0 & , \text{ sonst.} \end{cases} \end{aligned}$$

□

#### 3.2.3 Zufälliger Zeiteffekt

Bis jetzt haben die Modelle die gemeinsame Eigenschaft, dass die theoretische Querschnittkorrelation Null bleibt. Dies lässt sich dadurch begründen, dass der für die Querschnittkorrelation zuständige Faktor nicht im Modell aufgenommen wurde. Weil diese Korrelation in dieser Studie von besonderer Bedeutung ist, ist es von Interesse, den Daten die Möglichkeit zu geben, sich selbst über die Querschnittkorrelation zu äußern. Die Querschnittkorrelation bringt die Wirkung systematischer Risikoquellen zum Ausdruck, die nicht explizit im Modell enthalten sind. Dazu zählen zum Beispiel für den Zeitraum der Analyse die aus dem Bankenrettungspaket sowie der Insolvenzmeldung der US-amerikanischen Investmentbank Lehman Brothers entstehenden Risiken. Existieren ökonomische Faktoren, die sich auf die CDS-Preise betrachteter Unternehmen auswirken, so hat dies eine Korrelation zwischen CDS-Preisen zur Folge.

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Um unser Ziel zu verfolgen, wird angenommen, dass ein zeitabhängiger unbeobachtbarer zufälliger Effekt vorliegt ( $\sigma_\lambda^2 \neq 0$ ), der sich in den Reststörgrößen befindet. Das Modell hat die Form

$$v_{it} = \lambda_t + \varepsilon_{it}$$

Das gesamte Modell hat die Form

$$CDS = X\delta + (1_N \otimes I_T)\lambda + (I_N \otimes I_T)\varepsilon, \quad (83)$$

wobei

$$\lambda = (\lambda_1, \dots, \lambda_T)'$$

die Vektoren der zufälligen Effekte sind.

Auch hier sollen die Einflüsse der verzögerten endogenen Variable und die Variable iTraxx auf die CDS-Preise (im Modell (83)) analysiert werden. Dafür isolieren wir aufeinanderfolgend die Effekte dieser Variablen im Modell. Dabei setzen wir

- $\eta = 0$

$$X = \begin{pmatrix} 1_N \otimes 1_T & 1_N \otimes iTraxx \end{pmatrix} \in \mathbb{R}^{N \cdot T \times 2},$$

$$\delta = (\mu, \beta)' \in \mathbb{R}^2$$

für das *statische Modell mit zufälligem Zeiteffekt*. Wir beobachten, wie die Variable iTraxx ausschließlich die Ergebnisse der Modellierung mit zufälligen Unternehmens- und Zeiteffekten beeinflusst.

- $\eta \neq 0$  für das *dynamische Modell mit zufälligem Zeiteffekt*. Hier ist fraglich, wie groß der Einfluss der verzögerten CDS-Preise auf die Modellierung ist. Wir untersuchen dabei sowohl den Einfluss kombinierter Variablen im Modell, indem wir

- $\beta \neq 0$

$$X = \begin{pmatrix} 1_T & CDS_{1,-1} & iTraxx \\ 1_T & CDS_{2,-1} & iTraxx \\ \vdots & \vdots & \vdots \\ 1_T & CDS_{N,-1} & iTraxx \end{pmatrix} \in \mathbb{R}^{N \cdot T \times 3},$$

$$\delta = (\mu, \eta, \beta)' \in \mathbb{R}^3$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

setzen, als auch ausschließlich die Auswirkung der verzögerten CDS-Preise auf die Ergebnisse der Modellierung, indem wir

$$- \beta = 0$$

$$X = \begin{pmatrix} 1_T & CDS_{1,-1} \\ 1_T & CDS_{2,-1} \\ \vdots & \vdots \\ 1_T & CDS_{N,-1} \end{pmatrix} \in \mathbb{R}^{N \cdot T \times 2},$$

$$\delta = (\mu, \eta)' \in \mathbb{R}^2$$

setzen.

Unter Annahmen dieses Abschnittes lässt sich die Fehler-Kovarianzmatrix mithilfe der Formel (38) wie folgt umformen

$$\begin{aligned} \Omega &= \Omega(\lambda) + \Omega(\varepsilon) \\ &= \sigma_\varepsilon^2 \left[ \frac{\sigma_\lambda^2}{\sigma_\varepsilon^2} \Omega^*(\lambda) + \Omega^*(\varepsilon) \right] \\ &= \sigma_\varepsilon^2 [\kappa \Omega^*(\lambda) + \Omega^*(\varepsilon)] \\ &= \sigma_\varepsilon^2 \Omega^*, \end{aligned}$$

wobei

$$\kappa = \frac{\sigma_\lambda^2}{\sigma_\varepsilon^2}. \quad (84)$$

Die entsprechende REML-Funktion ist durch die folgende Form gegeben

$$\begin{aligned} l_R(\delta, \Omega) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\varepsilon^2 + \ln |X' \Omega^{*-1} X| + \ln |\Omega^*| \right] \\ &\quad - \frac{1}{2\sigma_\varepsilon^2} (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta). \end{aligned}$$

Unter Verwendung von

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N \cdot T - m} (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta), \quad (85)$$

erhält man die folgende restringierte Profile-Likelihood-Funktion

$$\begin{aligned} l_{Rp}(\delta, \kappa) &= -\frac{1}{2} (N \cdot T - m) \ln \left( (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta) \right) \\ &\quad - \frac{1}{2} \left[ \ln |X' \Omega^{*-1} X| + \ln |\Omega^*| \right] + const. \end{aligned}$$



### 3 Korrelation von Ausfallwahrscheinlichkeiten

Die Parameter  $\delta$  und  $\kappa$  lassen sich wieder iterativ durch Maximierung der  $l_{Rp}$ -Funktion bestimmen. Unter Ausnutzung der resultierenden Schätzer  $\hat{\delta}_{REML}$ ,  $\hat{\kappa}$  kann die Kovarianzmatrix  $\Omega^*$  geschätzt werden, sodass sich die Varianzparameter  $\hat{\sigma}_\varepsilon^2$ ,  $\hat{\sigma}_\lambda^2$  mittels (95) und (84) bestimmen lassen. Um die Korrelationsstruktur zu ermitteln, werden wir auch zwei Fälle betrachten, da die verzögerte exogene Variable abhängig von zufälligen Effekten ist.

#### Statisches Modell

Das Ziel ist hier ausschließlich den Einfluss der Variable  $iTraxx$  auf die Korrelation im Modell mit zufälligen Unternehmens- und Zeiteffekten zu ermitteln. Deshalb isolieren wir die verzögerten CDS-Preise im Modell ( $\eta = 0$ ). Das Modell sieht folgendermaßen aus

$$CDS_{it} = \mu + \beta iTraxx_t + \alpha_i + \lambda_t + \varepsilon_{it},$$

$$i = 1 \dots N, t = 1, \dots, T.$$

Die Korrelationsfunktion lässt sich wie folgt berechnen

$$\hat{\rho}_{ij}(t,s) = \begin{cases} 0 & , \text{ für } i = j, t \neq s, \\ \frac{\hat{\sigma}_\lambda^2}{\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2} & , \text{ für } i \neq j, t = s, \\ 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases}$$

Diese theoretische Korrelation deutet darauf hin, dass CDS-Preise nur über Banken ausschließlich am gleichen Tag korreliert sein können.

#### Dynamisches Modell

Hier untersuchen wir den Einfluss der verzögerten CDS-Preise auf die Korrelation bei der Modellierung mit zufälligen Unternehmens- und Zeiteffekten. Dabei setzen wir  $\eta \neq 0$ . Wir nutzen das folgende Modell aus

$$CDS_{it} = \eta^t CDS_{i,0} + \sum_{r=1}^t \eta^{t-r} (\mu + \beta iTraxx_r + v_{it}),$$

$$i = 1, \dots, N, t = 1, \dots, T. \quad (86)$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

**Lemma 3.5.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in dynamischen Modellen mit zufälligem Zeiteffekt durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \hat{\eta}^{|t-s|} \sqrt{\frac{(1-\hat{\eta}^{2\min(s, t)})}{(1-\hat{\eta}^{2\max(s, t)})}} & , \text{ für } i = j, \\ \frac{\hat{\sigma}_\lambda^2}{\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2} \hat{\eta}^{|t-s|} \sqrt{\frac{(1-\hat{\eta}^{2\min(s, t)})}{(1-\hat{\eta}^{2\max(s, t)})}} & , \text{ sonst.} \end{cases} \quad (87)$$

*Beweis.* Sei  $s, t \in \{1, 2, \dots, T\}$ , wir nehmen ohne Beschränkung der Allgemeinheit an, dass gilt  $t \geq s$ . Dann folgt

$$\begin{aligned} Kov_{ij}(t, s) &= kov(CDS_{it}, CDS_{js}) \\ &= kov\left(\sum_{r=1}^t \eta^{(t-r)} v_{ir}, \sum_{q=1}^s \eta^{(s-r)} v_{jq}\right) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\lambda_r + \varepsilon_{ir}, \lambda_q + \varepsilon_{jq}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} [kov(\lambda_r, \lambda_q) + kov(\varepsilon_{ir}, \varepsilon_{jq})] \\ &= \begin{cases} [\sigma_\varepsilon^2 + \sigma_\lambda^2] \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} & , \text{ für } i = j, \\ \sigma_\lambda^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} & , \text{ sonst.} \end{cases} \end{aligned}$$

Es folgt daraus, dass gilt

$$\rho_{ij}(t, s) = \begin{cases} \frac{\eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2}}{\sqrt{\frac{1-\eta^{2s}}{1-\eta^2}} \sqrt{\frac{1-\eta^{2t}}{1-\eta^2}}} & , \text{ für } i = j, \\ \frac{\sigma_\lambda^2}{[\sigma_\varepsilon^2 + \sigma_\lambda^2]} \frac{\eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2}}{\sqrt{\frac{1-\eta^{2s}}{1-\eta^2}} \sqrt{\frac{1-\eta^{2t}}{1-\eta^2}}} & , \text{ sonst.} \end{cases}$$

□

### AR(1)-Fehler

Unter der Annahme eines vorliegenden AR(1)-Prozesses hat der Fehlerterm die Form

$$\begin{aligned} v_{it} &= \alpha_i + \lambda_t + \varepsilon_{it}, \\ i &= 1, \dots, N, t = 1, \dots, T, \end{aligned}$$

$$\varepsilon_{it} = \phi \varepsilon_{i,t-1} + \zeta_{it}, \quad |\phi| < 1.$$

Die Kovarianzmatrix der Fehlerterme  $v_{it}$  mit vorliegender AR(1)-Prozess lautet

$$\begin{aligned} \Omega &= \sigma_\zeta^2 [\kappa^* \Omega^*(\lambda) + \mathbf{I}_N \otimes V(\phi)] \\ &= \sigma_\zeta^2 \Omega^{**}, \end{aligned} \tag{88}$$

wobei

$$\kappa^* = \frac{\sigma_\lambda^2}{\sigma_\zeta^2}, \tag{89}$$

Die entsprechende REML-Funktion hat die Form

$$\begin{aligned} l_R^{**}(\delta, \sigma_\zeta^2, \phi, d^*, \kappa^*) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\varepsilon^2 + \ln |X' \Omega^{**^{-1}} X| + \ln |\Omega^{**}| \right] \\ &\quad - \frac{1}{2\sigma_\zeta^2} (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta), \end{aligned} \tag{90}$$

Die Verteilungsannahmen über  $\zeta_{it}$  und die Notationen aus dem Abschnitt (3.2.2) bleiben erhalten. Unter Verwendung von

$$\hat{\sigma}_\zeta^2 = \frac{1}{N \cdot T - m} (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta), \tag{91}$$

in der REML-Funktion  $l_R^{**}$  (90), erhält man die folgende restringierte Profile-Likelihood-Funktion

$$\begin{aligned} l_{Rp}^{**}(\delta, \phi, d^*, \kappa^*) &= -\frac{1}{2} (N \cdot T - m) \ln \left( (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta) \right) \\ &\quad - \frac{1}{2} \left[ \ln |X' \Omega^{**^{-1}} X| + \ln |\Omega^{**}| \right] + const, \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

die mithilfe der Maximierungsalgorithmen maximiert werden kann und so die Modellparameter  $\delta, \phi, \kappa^*$  bestimmt. Setzt man die gefundenen Schätzer ( $\hat{\delta}_{REML}, \hat{\phi}, \hat{\kappa}^*$ ) in (97) ein, so kann der Kovarianzschätzer  $\hat{\Omega}^{**}$  bestimmt werden. Diese Kovarianzmatrix setzt man in die Gleichung (100) ein. Man erhält somit die Varianzschätzer  $\sigma_{REML}^2(\zeta)$ , der selber in die Gleichung (98) eingesetzt wird, sodass die Varianzkomponente  $\sigma_\lambda^2$  bestimmt werden kann.

Mithilfe der geschätzten Parameter lässt sich die Korrelationsfunktion berechnen. Wir erhalten folgende Funktion für das statische Modell mit zufälligen Unternehmens- und Zeiteffekten

**Lemma 3.6.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur im statischen Modell mit zufälligem Zeiteffekt und vorliegendem AR(1)-Prozess in Reststörgrößen durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t,s) = \begin{cases} \frac{\hat{\sigma}_\zeta^2 \hat{\phi}^{|t-s|}}{(1-\hat{\phi}^2)\hat{\sigma}_\lambda^2 + \hat{\sigma}_\zeta^2} & , \text{ für } i = j, t \neq s, \\ \frac{\hat{\sigma}_\lambda^2}{\hat{\sigma}_\lambda^2 + \frac{\hat{\sigma}_\zeta^2}{1-\hat{\phi}^2}} & , \text{ für } i \neq j, t = s, \\ 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases} \quad (92)$$

*Beweis.* Der Beweis folgt direkt aus der Kovarianzfunktion

$$\begin{aligned} kov(v_{it}, v_{js}) &= kov(\lambda_t, \lambda_s) + kov(\varepsilon_{it}, \varepsilon_{js}) \\ &= \begin{cases} \sigma_\zeta^2 \frac{\phi^{|t-s|}}{1-\phi^2} & , \text{ für } i = j, t \neq s, \\ \sigma_\lambda^2 + \frac{\sigma_\zeta^2}{1-\phi^2} & , \text{ für } i = j, t = s, \\ \sigma_\lambda^2 & , \text{ für } i \neq j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases} \end{aligned}$$

□

**Lemma 3.7.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in dynamischen Modellen mit zufälligem Zeiteffekt und vorliegendem AR(1)-Prozess in den Reststörgrößen durch die folgende Korrelationsfunktion

### 3 Korrelation von Ausfallwahrscheinlichkeiten

ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2\min(t,s)}}{1-\hat{\eta}^2} + \frac{\hat{\sigma}_\zeta^2}{1-\hat{\phi}^2} \sum_{r=1}^t \sum_{q=1}^s \hat{\eta}^{t-r+s-q} \hat{\phi}^{|r-q|}}{\sqrt{\hat{v}ar(CDS_{it})} \sqrt{\hat{v}ar(CDS_{js})}} & , \text{ für } i = j , \\ \frac{\hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2\min(t,s)}}{1-\hat{\eta}^2}}{\sqrt{\hat{v}ar(CDS_{it})} \sqrt{\hat{v}ar(CDS_{js})}} & , \text{ sonst.} \end{cases} .$$

Wobei

$$\hat{v}ar(CDS_{it}) = \hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2t}}{1-\hat{\eta}^2} + \hat{\sigma}_\zeta^2 \sum_{r=1}^t \sum_{q=1}^t \frac{\hat{\eta}^{2t-r-q} \hat{\phi}^{|r-q|}}{1-\hat{\phi}^2} .$$

**Beweis.** Wir nutzen die Formel (96) aus und erhalten die folgende Varianz-Kovarianzmatrix

$$\begin{aligned} Kov_{ij}(t, s) &= kov(CDS_{it}, CDS_{js}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\alpha_i + \lambda_r + \varepsilon_{ir}, \alpha_j + \lambda_q + \varepsilon_{jq}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} [kov(\alpha_i, \alpha_j) + kov(\lambda_r, \lambda_q) + kov(\varepsilon_{ir}, \varepsilon_{jq})] \end{aligned}$$

Damit ist für  $i = j$

$$\begin{aligned} Kov_{ii}(t, s) &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \sigma_\zeta^2 \frac{\phi^{|r-q|}}{1-\phi^2} + \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\lambda_r, \lambda_q) \\ &= \frac{\sigma_\zeta^2}{1-\phi^2} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \phi^{|r-q|} + \sigma_\lambda^2 \sum_{r=1}^{\min(t,s)} \eta^{2\{\min(t,s)-r\}} \\ &= \sigma_\lambda^2 \frac{1-\eta^{2\min(t,s)}}{1-\eta^2} + \frac{\sigma_\zeta^2}{1-\phi^2} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \phi^{|r-q|} \end{aligned}$$

und für  $i \neq j$

$$\begin{aligned} Kov_{ij}(t, s) &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\lambda_r, \lambda_q) \\ &= \sigma_\lambda^2 \sum_{r=1}^{\min(t,s)} \eta^{2\{\min(t,s)-r\}} \\ &= \sigma_\lambda^2 \frac{1-\eta^{2\min(t,s)}}{1-\eta^2} . \end{aligned}$$

□

### 3.2.4 Zufällige Unternehmens- und Zeiteffekte

Wir wollen hier untersuchen, wie die Korrelation auf die gleichzeitige Aufnahme von unternehmensspezifischen sowie zeitspezifischen Effekten im Modell reagiert. Wir betrachten deshalb das folgende Modell für die Reststörgrößen

$$v_{it} = \alpha_i + \lambda_t + \varepsilon_{it}$$

Der Anteil  $\alpha_i + \lambda_t$  wird als *zufällige Unternehmens- und Zeiteffekte* bezeichnet. Das gesamte Modell wird in der englischsprachigen Literatur als *two-way random-effects Model* geführt und hat die Form

$$CDS = X\delta + (I_N \otimes 1_T)\alpha + (1_N \otimes I_T)\lambda + (I_N \otimes I_T)\varepsilon, \quad (93)$$

wobei

$$\begin{aligned} \alpha &= (\alpha_1, \dots, \alpha_N)', \\ \lambda &= (\lambda_1, \dots, \lambda_T)' \end{aligned}$$

die Vektoren der zufälligen Effekte sind.

Auch hier sollen die Einflüsse der verzögerten endogenen Variable und die Variable *iTraxx* auf die CDS-Preise (im Modell (93)) analysiert werden. Die entsprechenden Designmatrizen ( $X$ ) sind im Abschnitt 3.2.3 definiert. Unter Annahmen dieses Abschnittes lässt sich die Fehler-Kovarianzmatrix mithilfe der Formel (38) wie folgt umformen

$$\begin{aligned} \Omega &= \Omega(\alpha) + \Omega(\lambda) + \Omega(\varepsilon) \\ &= \sigma_\varepsilon^2 \left[ \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} \Omega^*(\alpha) + \frac{\sigma_\lambda^2}{\sigma_\varepsilon^2} \Omega^*(\lambda) + \Omega^*(\varepsilon) \right] \\ &= \sigma_\varepsilon^2 [d\Omega^*(\alpha) + \kappa\Omega^*(\lambda) + \Omega^*(\varepsilon)] \\ &= \sigma_\varepsilon^2 \Omega^*, \end{aligned}$$

wobei

$$d = \frac{\sigma_\alpha^2}{\sigma_\varepsilon^2} \quad \text{und} \quad \kappa = \frac{\sigma_\lambda^2}{\sigma_\varepsilon^2}. \quad (94)$$

Die entsprechende REML-Funktion ist durch die folgende Form gegeben

$$\begin{aligned} l_R(\delta, \Omega) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\varepsilon^2 + \ln |X' \Omega^{*-1} X| + \ln |\Omega^*| \right] \\ &\quad - \frac{1}{2\sigma_\varepsilon^2} (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta). \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Unter Verwendung von

$$\hat{\sigma}^2(\varepsilon) = \frac{1}{N \cdot T - m} (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta), \quad (95)$$

erhält man die folgende restringierte Profile-Likelihood-Funktion

$$l_{Rp}(\delta, d, \kappa) = -\frac{1}{2}(N \cdot T - m) \ln \left( (CDS - X\delta)' \Omega^{*-1} (CDS - X\delta) \right) \\ - \frac{1}{2} \left[ \ln |X' \Omega^{*-1} X| + \ln |\Omega^*| \right] + const.$$

Die Parameter  $\delta$ ,  $d$  und  $\kappa$  lassen sich wieder iterativ durch Maximierung der  $l_{Rp}$ -Funktion bestimmen. Unter Ausnutzung der resultierenden Schätzer  $\hat{\delta}$ ,  $\hat{d}$ ,  $\hat{\kappa}$  kann die Kovarianzmatrix  $\Omega^*$  geschätzt werden, sodass sich die Varianzparameter  $\hat{\sigma}_{REML}^2(\varepsilon)$ ,  $\hat{\sigma}_{REML}^2(\alpha)$ ,  $\hat{\sigma}_{REML}^2(\lambda)$  mittels (95) und (98) bestimmen lassen.

Um die Korrelationsstruktur zu ermitteln, werden wir wieder zwei Fälle betrachten, da die verzögerte exogene Variable abhängig von zufälligen Effekten ist.

#### Statisches Modell

Das Ziel ist hier ausschließlich den Einfluss der Variable *iTraxx* auf die Korrelation im Modell mit zufälligen Unternehmens- und Zeiteffekten zu ermitteln. Deshalb isolieren wir die verzögerten CDS-Preise im Modell ( $\eta = 0$ ). Das Modell sieht folgendermaßen aus

$$CDS_{it} = \mu + \beta iTraxx_t + \alpha_i + \lambda_t + \varepsilon_{it}, \\ i = 1 \dots N, t = 1, \dots, T.$$

Die Korrelationsfunktion lässt sich wie folgt berechnen (vergleiche Baltagi, 2003, S. 34)

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_{\alpha}^2}{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\lambda}^2} & , \text{ für } i = j, t \neq s, \\ \frac{\hat{\sigma}_{\lambda}^2}{\hat{\sigma}_{\alpha}^2 + \hat{\sigma}_{\varepsilon}^2 + \hat{\sigma}_{\lambda}^2} & , \text{ für } i \neq j, t = s, \\ 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases}$$

Diese theoretische Korrelation deutet darauf hin, dass CDS-Preise über Banken ausschließlich am gleichen Tag korreliert sein können. Außerdem sind CDS-Preise konstant über die Zeit korreliert.

### 3 Korrelation von Ausfallwahrscheinlichkeiten

#### Dynamisches Modell

Hier untersuchen wir den Einfluss der verzögerten CDS-Preise auf die Korrelation bei der Modellierung mit zufälligen Unternehmens- und Zeiteffekten. Dabei setzen wir  $\eta \neq 0$ . Wir nutzen das folgende Modell aus:

$$CDS_{it} = \eta^t CDS_{i,0} + \sum_{r=1}^t \eta^{t-r} (\mu + \beta i Traxx_r + v_{it}),$$

$$i = 1, \dots, N, t = 1, \dots, T. \quad (96)$$

**Lemma 3.8.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in dynamischen Modellen mit zufälligen Unternehmens- und Zeiteffekten durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_\alpha^2 \frac{1-\eta^t}{1-\eta} \frac{1-\eta^s}{1-\eta} + [\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2] \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2}}{\sqrt{\hat{\sigma}_\alpha^2 \left(\frac{1-\eta^s}{1-\eta}\right)^2 + [\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2] \frac{1-\eta^{2s}}{1-\eta^2}} \sqrt{\hat{\sigma}_\alpha^2 \left(\frac{1-\eta^t}{1-\eta}\right)^2 + [\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2] \frac{1-\eta^{2t}}{1-\eta^2}}} , & \text{für } i = j, \\ \frac{\hat{\sigma}_\lambda^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2}}{\sqrt{\hat{\sigma}_\alpha^2 \left(\frac{1-\eta^s}{1-\eta}\right)^2 + [\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2] \frac{1-\eta^{2s}}{1-\eta^2}} \sqrt{\hat{\sigma}_\alpha^2 \left(\frac{1-\eta^t}{1-\eta}\right)^2 + [\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_\lambda^2] \frac{1-\eta^{2t}}{1-\eta^2}}} , & \text{sonst.} \end{cases}$$

*Beweis.* Sei  $s, t \in \{1, 2, \dots, T\}$ , wir nehmen ohne Beschränkung der Allgemeinheit an, dass gilt  $t \geq s$ . Dann folgt:

$$\begin{aligned} Kov_{ij}(t, s) &= kov(CDS_{it}, CDS_{js}) \\ &= kov\left(\sum_{r=1}^t \eta^{(t-r)} v_{ir}, \sum_{q=1}^s \eta^{(s-r)} v_{jq}\right) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\alpha_i + \lambda_r + \varepsilon_{ir}, \alpha_j + \lambda_q + \varepsilon_{jq}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} [kov(\alpha_i, \alpha_j) + kov(\lambda_r, \lambda_q) + kov(\varepsilon_{ir}, \varepsilon_{jq})] \\ &= \begin{cases} \sigma_\alpha^2 \frac{1-\eta^t}{1-\eta} \frac{1-\eta^s}{1-\eta} + [\sigma_\varepsilon^2 + \sigma_\lambda^2] \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} , & \text{für } i = j, \\ \sigma_\lambda^2 \eta^{t-s} \frac{1-\eta^{2s}}{1-\eta^2} , & \text{sonst.} \end{cases} \end{aligned}$$

□



### AR(1)-Fehler

Unter der Annahme eines vorliegenden AR(1)-Prozesses hat der Fehlerterm die Form

$$v_{it} = \alpha_i + \lambda_t + \varepsilon_{it},$$

$$i = 1, \dots, N, t = 1, \dots, T,$$

$$\varepsilon_{it} = \phi \varepsilon_{i,t-1} + \zeta_{it}, \quad |\phi| < 1.$$

Die Verteilungsannahmen über  $\zeta_{it}$  und die Notationen aus dem Abschnitt (3.2.2) bleiben erhalten. Die Kovarianzmatrix des Fehlerterms  $v_{it}$  mit vorliegender AR(1)-Prozess lautet

$$\begin{aligned} \Omega &= \sigma_\zeta^2 [d^* \Omega^*(\alpha) + \kappa^* \Omega^*(\lambda) + I_N \otimes \text{Korr}] \\ &= \sigma_\zeta^2 \Omega^{**}, \end{aligned} \quad (97)$$

wobei

$$d^* = \frac{\sigma_\alpha^2}{\sigma_\zeta^2} \quad \text{und} \quad \kappa^* = \frac{\sigma_\lambda^2}{\sigma_\zeta^2}. \quad (98)$$

Die entsprechende REML-Funktion hat die Form

$$\begin{aligned} l_R^{**}(\delta, \sigma_\zeta^2, \phi, d^*, \kappa^*) &= -\frac{1}{2} \left[ (N \cdot T - m) \ln \sigma_\varepsilon^2 + \ln |X' \Omega^{**^{-1}} X| + \ln |\Omega^{**}| \right] \\ &\quad - \frac{1}{2\sigma_\zeta^2} (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta), \end{aligned} \quad (99)$$

unter Verwendung von

$$\hat{\sigma}^2(\zeta) = \frac{1}{N \cdot T - m} (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta), \quad (100)$$

in der REML-Funktion  $l_R^{**}$  (99), erhält man die folgende restringierte Profile-Likelihood-Funktion

$$\begin{aligned} l_{Rp}^{**}(\delta, \phi, d^*, \kappa^*) &= -\frac{1}{2} (N \cdot T - m) \ln \left( (CDS - X\delta)' \Omega^{**^{-1}} (CDS - X\delta) \right) \\ &\quad - \frac{1}{2} \left[ \ln |X' \Omega^{**^{-1}} X| + \ln |\Omega^{**}| \right] + \text{const}, \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

die mithilfe der Maximierungsalgorithmen maximiert werden kann und so die Modellparameter  $\delta, \phi, d^*, \kappa^*$  bestimmt. Setzt man die gefundenen Schätzer ( $\hat{\delta}_{REML}, \hat{\phi}, \hat{d}^*, \hat{\kappa}^*$ ) in (97) ein, so kann der Kovarianzschätzer  $\hat{\Omega}^{**}$  bestimmt werden. Diese Kovarianzmatrix setzt man in die Gleichung (100) ein. Man erhält somit die Varianzschätzer  $\sigma_{REML}^2(\zeta)$ , der selber in die Gleichung (98) eingesetzt wird, sodass die anderen Varianzkomponenten  $\sigma_{REML}^2(\alpha), \sigma_{REML}^2(\lambda)$  bestimmt werden können. Mithilfe der geschätzten Parameter lässt sich die Korrelationsfunktion berechnen.

**Lemma 3.9.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in statischem Modell mit zufälligen Unternehmens- und Zeiteffekten sowie vorliegendem  $AR(1)$ -Prozess in Reststörgrößen durch die folgende Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\zeta^2 \frac{\hat{\phi}^{|t-s|}}{1-\hat{\phi}^2}}{\hat{\sigma}_\zeta^2} & , \text{ für } i = j, t \neq s, \\ \frac{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\lambda^2 + \frac{\hat{\sigma}_\zeta^2}{1-\hat{\phi}^2}}{\hat{\sigma}_\lambda^2} & , \text{ für } i \neq j, t = s, \\ 1 & , \text{ für } i = j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases} \quad (101)$$

*Beweis.* Dies lässt sich einfach beweisen, indem man die Kovarianz wie folgt schreibt

$$\begin{aligned} kov(v_{it}, v_{js}) &= kov(\alpha_i, \alpha_j) + kov(\lambda_t, \lambda_s) + kov(\varepsilon_{it}, \varepsilon_{js}) \\ &= \begin{cases} \sigma_\alpha^2 + \sigma_\zeta^2 \frac{\phi^{|t-s|}}{1-\phi^2} & , \text{ für } i = j, t \neq s, \\ \sigma_\alpha^2 + \sigma_\lambda^2 + \frac{\sigma_\zeta^2}{1-\phi^2} & , \text{ für } i = j, t = s, \\ \sigma_\lambda^2 & , \text{ für } i \neq j, t = s, \\ 0 & , \text{ für } j \neq i, t \neq s. \end{cases} \end{aligned}$$

□

**Lemma 3.10.** Für zwei beliebige Unternehmen  $i, j$  und zwei diskrete Zeitpunkte  $t, s$  lässt sich die Abhängigkeitsstruktur in dynamischen Modellen mit zufälligen Unternehmens- und Zeiteffekten sowie vorliegendem  $AR(1)$ -Prozess in Reststörgrößen durch die folgende

### 3 Korrelation von Ausfallwahrscheinlichkeiten

Korrelationsfunktion ermitteln:

$$\hat{\rho}_{ij}(t, s) = \begin{cases} \frac{\frac{\hat{\sigma}_\alpha^2(1-\hat{\eta}^t)(1-\hat{\eta}^s)}{(1-\hat{\eta})^2} + \hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2\min(t, s)}}{1-\hat{\eta}^2} + \frac{\hat{\sigma}_\zeta^2}{1-\hat{\phi}^2} \sum_{r=1}^t \sum_{q=1}^s \hat{\eta}^{t-r+s-q} \hat{\phi}^{|r-q|}}{\sqrt{\hat{v}ar(CDS_{it})} \sqrt{\hat{v}ar(CDS_{js})}} & , \text{ für } i = j, \\ \frac{\hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2\min(t, s)}}{1-\hat{\eta}^2}}{\sqrt{\hat{v}ar(CDS_{it})} \sqrt{\hat{v}ar(CDS_{js})}} & , \text{ sonst.} \end{cases}$$

Wobei

$$\hat{v}ar(CDS_{it}) = \hat{\sigma}_\alpha^2 \left( \frac{1-\hat{\eta}^t}{1-\hat{\eta}} \right)^2 + \hat{\sigma}_\lambda^2 \frac{1-\hat{\eta}^{2t}}{1-\hat{\eta}^2} + \hat{\sigma}_\zeta^2 \sum_{r=1}^t \sum_{q=1}^t \frac{\hat{\eta}^{2t-r-q} \hat{\phi}^{|r-q|}}{1-\hat{\phi}^2}.$$

**Beweis.** Wir nutzen (96) aus und erhalten die folgende Kovarianzfunktion:

$$\begin{aligned} Kov_{ij}(t, s) &= kov(CDS_{it}, CDS_{js}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\alpha_i + \lambda_r + \varepsilon_{ir}, \alpha_j + \lambda_q + \varepsilon_{jq}) \\ &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} [kov(\alpha_i, \alpha_j) + kov(\lambda_r, \lambda_q) + kov(\varepsilon_{ir}, \varepsilon_{jq})]. \end{aligned}$$

Damit ist für  $i = j$ :

$$\begin{aligned} Kov_{ii}(t, s) &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \left[ \sigma_\alpha^2 + \sigma_\zeta^2 \frac{\phi^{|r-q|}}{1-\phi^2} \right] \\ &\quad + \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} kov(\lambda_r, \lambda_q) \\ &= \frac{\sigma_\alpha^2(1-\eta^t)(1-\eta^s)}{(1-\eta)^2} + \frac{\sigma_\zeta^2}{1-\phi^2} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \phi^{|r-q|} \\ &\quad + \sigma_\lambda^2 \sum_{r=1}^{\min(t, s)} \eta^{2\{\min(t, s)-r\}} \\ &= \frac{\sigma_\alpha^2(1-\eta^t)(1-\eta^s)}{(1-\eta)^2} + \sigma_\lambda^2 \frac{1-\eta^{2\min(t, s)}}{1-\eta^2} \\ &\quad + \frac{\sigma_\zeta^2}{1-\phi^2} \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \phi^{|r-q|} \end{aligned}$$

### 3 Korrelation von Ausfallwahrscheinlichkeiten

und für  $i \neq j$ :

$$\begin{aligned} \text{Kov}_{ij}(t, s) &= \sum_{r=1}^t \sum_{q=1}^s \eta^{t-r+s-q} \text{kov}(\lambda_r, \lambda_q) \\ &= \sigma_\lambda^2 \sum_{r=1}^{\min(t, s)} \eta^{2\{\min(t, s)-r\}} \\ &= \sigma_\lambda^2 \frac{1 - \eta^{2\min(t, s)}}{1 - \eta^2}. \end{aligned}$$

□

## 4 Material und Auswertung

### 4.1 Material

Gemessen an der Bilanzsumme zum 31. Dezember 2008 sind die 20 größten deutschen Kreditinstitute (ohne die Deutsche Bundesbank) die Deutsche Bank AG, die Commerzbank AG, die bayerische Hypo- und Vereinsbank AG (HypoVereinsbank), die Landesbank Baden-Württemberg AöR (LBBW), die DZ Bank AG, die Bayerische Landesbank AöR (BayernLB), die Dresdner Bank AG, die Hypo Real Estate Holding AG (HRE), die KfW, die WestLB AG, die Eurohypo AG, die Norddeutsche Landesbank Girozentrale AöR (NordLB), die Postbank AG, die HSH Nordbank AG, die Landesbank Hessen-Thüringen Girozentrale AöR (HeLaBa), die NRW.Bank AöR, die Landesbank Berlin Holding AG (LB Berlin), die DekaBank Deutsche Girozentrale AöR (Deka Bank), die Westdeutsche Genossenschafts-Zentralbank Bank AG (WGZ) und die Landwirtschaftliche Rentenbank AöR (Rentenbank). Gemeinsam weisen diese 20 Banken einen Anteil von 82,1 % beziehungsweise 80,1 % für das Jahr 2008 beziehungsweise 2007 an den – über die 100 führenden deutschen Kreditinstitute – aggregierten Bilanzsummen auf. Eine Übersicht der von uns betrachteten Banken findet sich in Tabelle 1.

Zur HRE, KfW, NRW.Bank, Deka Bank, Eurohypo, WGZ und Rentenbank waren keine CDS-Preise verfügbar. Wir haben sie deshalb aus unserer Analyse herausgenommen und uns auf die verbleibenden 13 Banken konzentriert. Die Dresdner Bank wurde von der Commerzbank gekauft. Trotzdem sind beide Banken in der Tabelle 1 noch getrennt gelistet, da die Übernahme erst 2009 endgültig vollzogen wurde. Diese 13 Banken weisen selber gemeinsam einen Anteil von 80,3 % beziehungsweise 81,2 % für das Jahr 2008 beziehungsweise 2007 an den – über die 20 größten führenden deutschen Kreditinstitute – aggregierten Bilanzsummen der Top-20-Kreditinstitute auf, sodass wir davon ausgehen können, dass die in den vorherigen Abschnitten untersuchte Abhängigkeitsmodellierung der CDS-Preise nicht nur die genannten 13 Banken sondern auch allgemein für die größten deutschen Kreditinstitute gültig ist.

Als CDS-Prämien liegen uns die Fünf-Jahres-CDS-Prämien der 13 in Tabelle 1 genannten Banken vom 2. Juli 2007 bis zum 7. Januar 2009 vor. Die Daten haben wir über Bloomberg bezogen. Insgesamt 394 Handelstage ist die Fensterlänge der Studie. Es lagen uns nur für die Dresdner Bank vollständige Daten vor. Für alle anderen Banken waren die Daten nicht vollständig, es fehlten mindestens neun Handelstage pro Bank. Bei der Postbank liegen erst Beobachtungen ab dem 8. Oktober 2007 vor, also fehlen insgesamt 88 Handelstage.

#### 4 Material und Auswertung

Tabelle 1: Bilanzsummen der betrachteten Banken.

Bank	Bilanzsumme in Milliarden €	
	zum 31. Dez. 2008	zum 31. Dez. 2007
Deutsche Bank	2.202,423	1.925,003
Commerzbank	625,196	616,196
HypoVereinsbank	458,602	422,129
LBBW	447,932	443,400
DZ Bank	427,090	431,337
BayernLB	421,666	415,639
Dresdner Bank	420,961	500,209
WestLB	288,122	287,416
NordLB	244,265	201,540
Postbank	231,282	202,913
HSH Nordbank	208,401	204,827
HeLaBa	184,572	173,787
LB Berlin	145,371	142,163

Quelle: Kuck (2009).

Für die Variable iTraxx fehlen uns insgesamt Beobachtungen für neun Handelstage. Damit wir gute Eigenschaften für die Schätzer haben und somit bessere Ergebnisse erzielen, modellieren wir die Zielvariable (CDS) durch ein balanciertes Modell. Diese Art von Modell erfordert, dass alle Banken die gleiche Beobachtungslänge haben müssen. Wir haben folgende Möglichkeiten:

- Wir wählen entweder nur Handelstage, für die Beobachtungen für alle Banken sowie für iTraxx vorliegen,
- oder wir wählen 394 Handelstage als Fensterlänge und ersetzen die fehlenden Daten durch prognostizierte Werte.

Die erste Möglichkeit scheint einfach und realistisch zu sein. Da wir aber auch großen Wert auf die Korrelation über die Zeit legen, wird dieser Ansatz nicht optimal sein, da hinter zwei aufeinanderfolgenden Tagen in den Daten sich viele Handelstage in der Realität verstecken könnten. Eine Folge davon ist, dass die Korrelation nicht mehr sinnvoll ermittelt und interpretiert werden kann. Hinzu kommt, dass durch die Streichung von Handelstagen, für die Daten für einige aber nicht alle Banken vorliegen, zu viele Informationen verloren gehen.

#### 4 Material und Auswertung

Bei der zweiten Alternative ist mit einigen Vorhersagefehlern zu rechnen. Aber wir hoffen, ein gutes erklärendes Modell könnte dazu beitragen, große Vorhersagefehler innerhalb des Beobachtungszeitraumes zu vermeiden, sodass plausible Ergebnisse bei der Modellierung erzielt werden können. Um die fehlenden Werte bei dem CDS sowie bei dem iTraxx schätzen zu können, haben wir ein multilineares Modell ausgenutzt. Insgesamt wurde das niedrigste  $R_{adj}^2 = 0,91$  bei der Postbank erreicht. Die anderen Banken weisen eine  $R_{adj}^2$  größer 0,95 auf. Die Anpassung des Modells beim iTraxx erfolgt mit einem  $R_{adj}^2 = 0,99$ , sodass wir sagen können, die aus Vorhersagen entstehenden Fehler sind in diesem Modell sehr niedrig.

Bevor wir mit der Auswertung beginnen, stellen wir in Abbildung 3 die beiden Variablen dar. Die Grafik lässt einen starken Zusammenhang zwischen beiden Variablen vermuten, da sie gleichzeitig zusammen steigen und sinken. In Abbildung 3 ist außerdem zu erkennen, dass die CDS-Preise bei der Postbank keine große Schwankungen aufweisen, dementsprechend vorsichtig muss die folgende Modellanpassung für diese Bank interpretiert werden.

Der Markit-iTraxx-Europe-Index repräsentiert das mittlere Risikoniveau, das auf dem Markt in den Regionen Europa und Asien gehandelt wird. Liegt der Index eines Unternehmens unter dem Marktindex, so hat das Unternehmen eher ein gutes Verhalten im Vergleich zu der allgemeinen Bewertung der am Kreditderivate-Markt aktiv agierenden Teilnehmern. In der Abbildung 4 ist der standardisierte Abstand zwischen dem Markit-iTraxx-Europe-Index und den Unternehmenindizes dargestellt, damit sich ein Überblick über die betrachteten Banken auf den Markt verschafft werden kann. Außer der HSH Nordbank und der damaligen WestLB haben sich die betrachteten Banken im Allgemein gut auf dem Markt vom 2. Juli 2007 bis kurz vor dem 15. September 2008 verhalten. Am 15. September 2008 erfolgte die Insolvenzmeldung der US-amerikanischen Investmentbank Lehman Brothers. Diese Entscheidung hat sich in Schockwellen rund um den Globus verbreitet: Die Aktienkurse stürzten ab, der US-Leitindex Dow Jones erlitt den stärksten Tagesverlust seit den Terrorattacken vom 11. September 2001, die CDS-Indices von Deutschen Kreditinstituten wiesen einen außerordentlichen Sprung nach oben auf (vergleiche Abbildung 4). In Deutschland machten insbesondere die Landesbanken, darunter die HSN NordBank, die BayernLB sowie die damalige West LB Milliardenverluste. Dies ist insbesondere stark bei der HSN NordBank. Der Hintergrund der Verluste bei der HSN NordBank ist ein Geschäft mit CDS: Die US-amerikanische Bank Goldman Sachs hatte sich bei der HSH NordBank gegen Kreditausfälle von Lehman Brothers versichert. Nach der Pleite des Hauses im September 2008 hat die Landesbank die Zahlung in Höhe von 45 Millionen US-Dollar – wie aus der Abbildung 2

4 Material und Auswertung

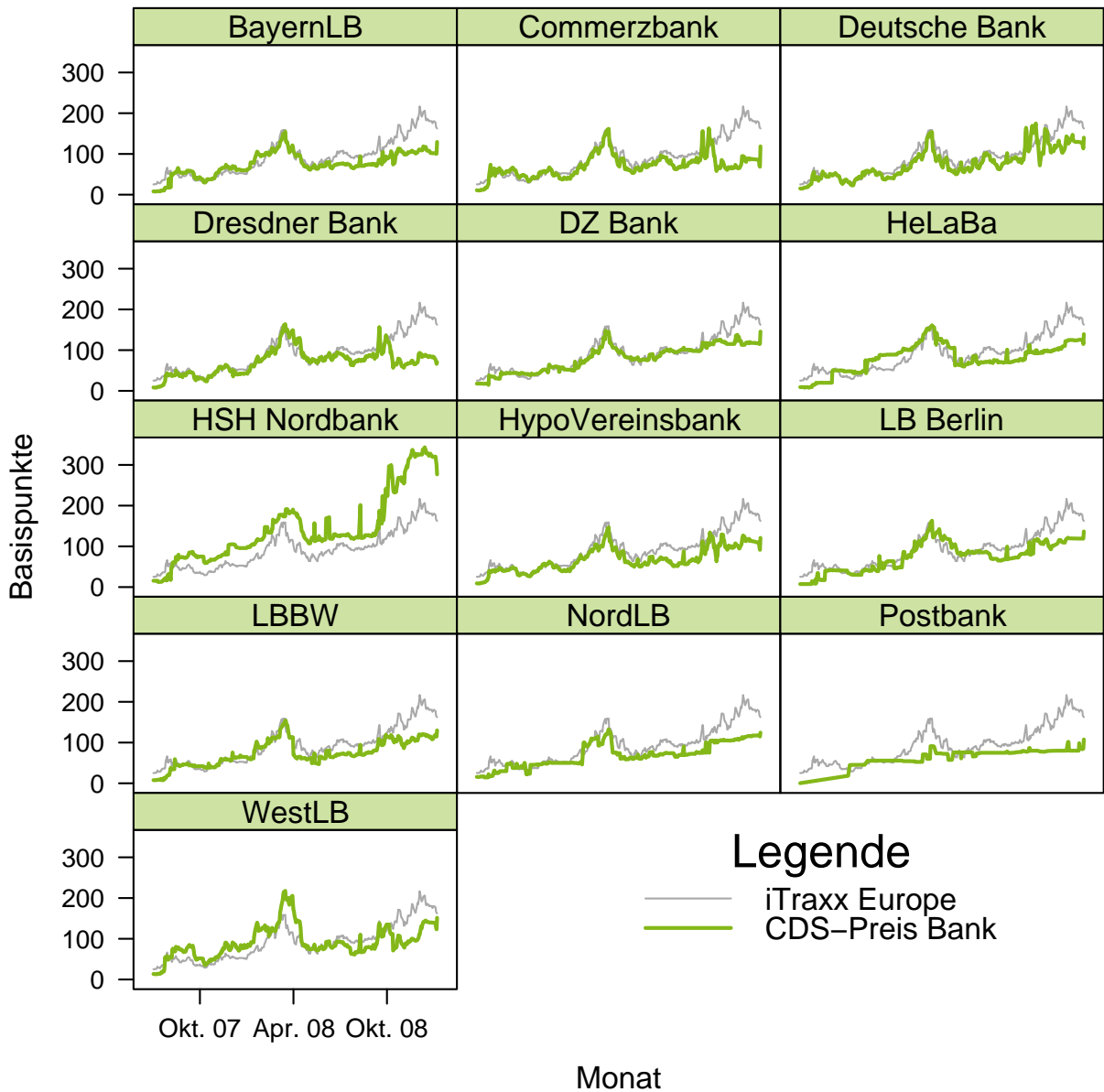


Abbildung 3: Zeitreihen der Fünf-Jahres-CDS-Preise der 13 Banken (grün) und des Markt iTraxx Europe (grau) vom 2. Juli 2007 bis zum 7. Januar 2009.



#### 4 Material und Auswertung

zu sehen ist – ausgeglichen (vergleiche Die Zeit (2009)). Um die Finanzmärkte zu

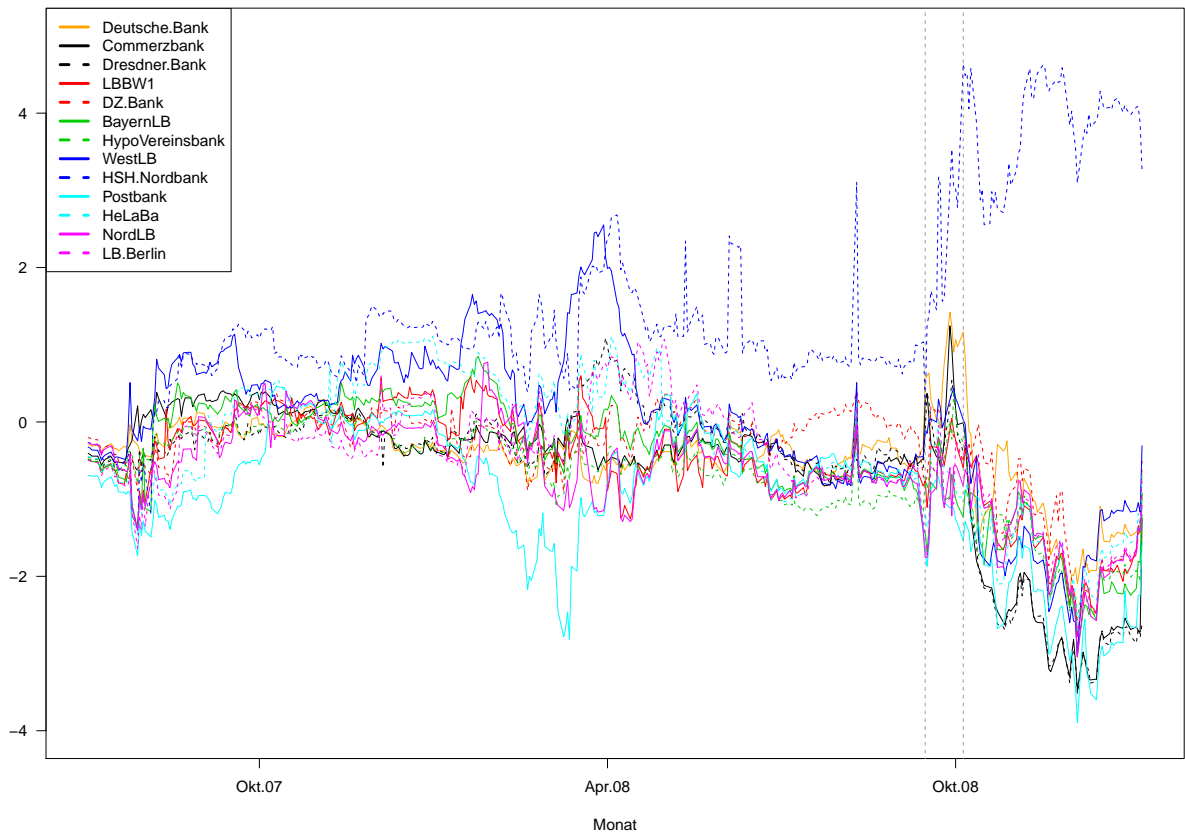


Abbildung 4: Zeitreihen des standardisierten Abstandes zwischen den Fünf-Jahres-CDS-Preisen und Markt-iTraxx-Europe-Indices der 13 Banken vom 2. Juli 2007 bis zum 7. Januar 2009 ( $\frac{\text{CDS}-\text{iTraxx}}{\sigma}$ ). Die erste und zweite vertikale gestrichelte graue Linie (vor und nach Oktober 2008) entsprechen dem 15. September beziehungsweise dem 5. Oktober 2008

beruhigen und die Krise einzudämmen, begannen nach und nach die Regierungen weltweit Milliardenbeträge – ein Großteil davon als Bürgschaften – zur Rettung der in Not geratenen Banken bereitzustellen. Am Sonntag, dem 5. Oktober 2008 traten Bundeskanzlerin Angela Merkel und der damalige Bundesfinanzminister Peer Steinbrück vor die Presse und gaben eine Garantie für sämtliche Spareinlagen in Deutschland. Dies stellte insofern eine Wende in der Regierungspolitik dar, als dass erstmals eine Maßnahme angekündigt wurde, die den gesamten Bankensektor

#### 4 Material und Auswertung

betrif. Zuvor hatte die Bundesregierung stets ein solches Vorgehen abgelehnt und auf Einzelfalllösungen wie bei der HRE bestanden. Bereits am darauffolgenden Handelstag kamen erste Gerüchte über einen Plan „B“ auf, der einen Rettungsschirm für den gesamten deutschen Bankensektor umfassen sollte. In den folgenden Tagen wurden in der Öffentlichkeit verschiedene Rettungspläne diskutiert, etwa ein gemeinsamer Rettungsschirm aller Mitgliedstaaten der Europäischen Union, eine Teilverstaatlichung nach britischem Vorbild oder Staatsgarantien für die betroffenen Banken. Am 11. Oktober wurden schließlich erste Pläne für ein Eilgesetz bekannt und bereits zwei Tage später verabschiedete das Bundeskabinett das FMStG und das Gesetz zur Errichtung eines Finanzmarktstabilisierungsfonds (Finanzmarktstabilisierungsfondsgesetz; FMStFG). Am 15. Oktober fand die erste Lesung im Bundestag statt, am Freitag, dem 17. Oktober wurde das Gesetz im Bundestag verabschiedet. Noch am selben Tag stimmte auch der Bundesrat zu und der damalige Bundespräsident Horst Köhler unterzeichnete das Gesetz. Mit der Verordnung zur Durchführung des Finanzmarktstabilisierungsfondsgesetzes (FMStFV) vom 20. Oktober wurde der von der neu geschaffenen Finanzmarktstabilisierungsanstalt verwaltete Finanzmarktstabilisierungsfonds (SoFFin) eingerichtet. Bereits zwei Tage später nahm die BayernLB als erste Bank Mittel aus dem SoFFin in Anspruch. Zwar erregte diese Inanspruchnahme noch einmal ein großes öffentliches Interesse, der Fokus der Berichterstattung wechselte jedoch auf das Konjunkturpaket, das bereits am 5. November 2008 verabschiedet wurde.

Allein das Hilfspaket der Deutsche Regierung umfasste 500 Milliarden Euro, die aus Bürgschaften für Interbanken-Kredite in Höhe von 400 Milliarden sowie weitere 100 Milliarden für Aufnahme an Krediten bestanden. Ein langfristiger Effekt dieser Staatsinterventionen ist schon ab den 5. Oktober in der Abbildung 4 zu erkennen: der standardisierte Abstand zwischen Marktindex und betrachteten Unternehmensindizes springt diesmal nach unten und Ende Oktober liegt er sogar unter Null. Das heißt, dass die Unternehmensindices der betrachteten Banken deutlich unter dem Marktindex liegen. Trotz dieser staatlichen Maßnahmen fuhr die HSN NordBank 2008 wegen massiver Abschreibungen in Folge der Finanzkrise einen Verlust von 2,8 Milliarden Euro ein und stand damit kurz vor der Pleite. Nur durch ein milliardenschweres Rettungspaket der Länder Hamburg und Schleswig-Holstein konnte die Bank gerettet werden. Insgesamt ist festzustellen, dass ohne systematische Risiken ein bedeutender Zusammenhang zwischen iTraxx und CDS-Preisen der betrachteten Banken besteht.

Berechnet man die prozentualen Änderungen in CDS-Preisen bei den betrachteten Banken, so weisen alle Banken eine schwache Änderung auf, die maximale

#### 4 Material und Auswertung

Änderung zwischen aufeinanderfolgenden CDS-Preisen liegt bei 3 %. In der Abbildung 5 sind diese Änderungen dargestellt. Die sehr niedrige Änderungen zwischen aufeinanderfolgenden Preisen weisen auf eine höhere Autokorrelation hin, was an ein dynamisches Modell für die Modellierung denken lässt.

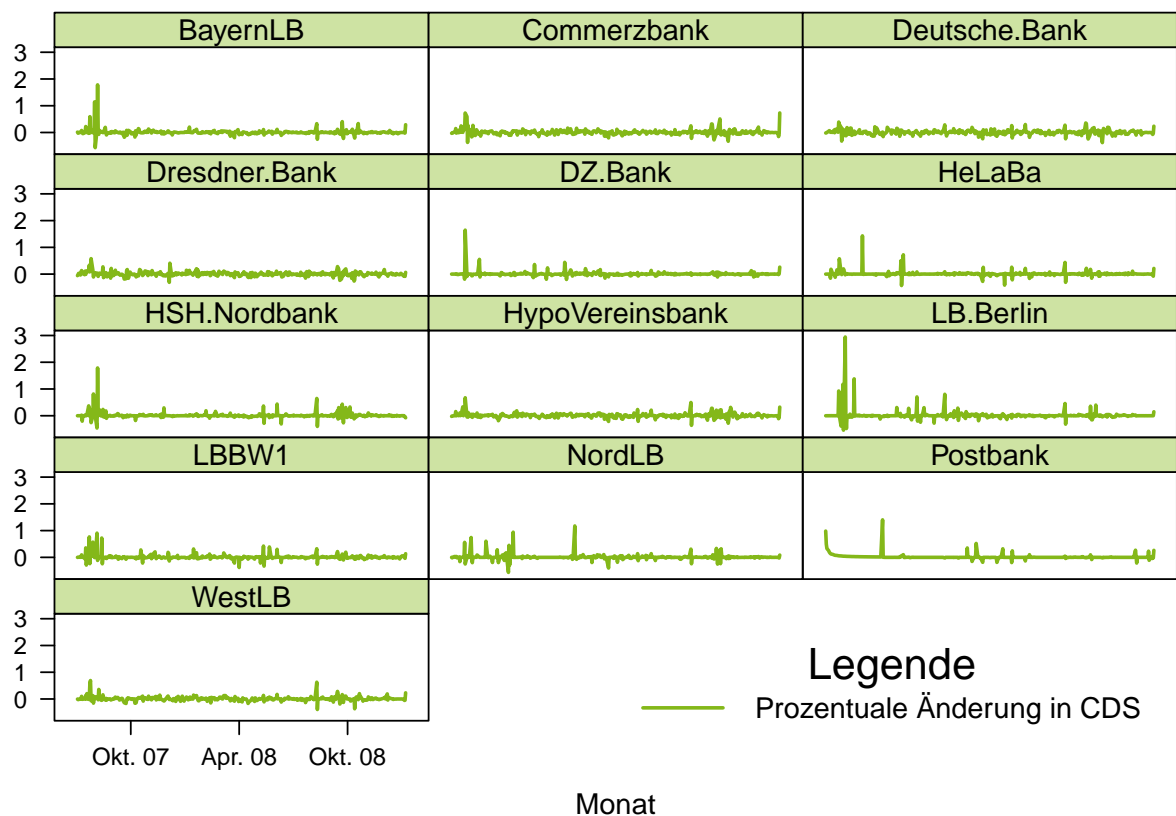


Abbildung 5: Zeitreihen der prozentualen Änderung in Fünf-Jahres-CDS-Preisen der 13 Banken vom 2. Juli 2007 bis zum 7. Januar 2009.

Berechnet man jetzt die Längsschnittkorrelation, so weisen alle Banken den höchsten Koeffizienten bei einem Zeitabstand von 1 auf, er liegt stets über 0,95. Bei der Querschnittkorrelation liegt dieser Koeffizient dagegen zwischen 0,52 und 0,95. Dementsprechend lehnt der einseitige Durbin-Watson-Test (vergleiche Zeileis und Hothorn (2002)) die Nullhypothese, die Längsschnittkorrelation betrage 0, zum 5 %-

#### *4 Material und Auswertung*

Niveau für alle Banken ab. Die Tabelle 2 zeigt, dass die CDS-Preise für alle Banken über die Zeit hoch korreliert sind. Die berechneten Pearson-Autokorrelationen geben einen Hinweis darüber an, wie groß die geschätzten Längsschnittkorrelationen bei der Modellierung zu erwarten sind. Wie Tabelle 3 zu entnehmen ist, gilt dies auch für die Querschnittkorrelation. Insgesamt beträgt die gemittelte paarweise Korrelation zwischen den Banken 0,80. Ein 5 %-Konfidenzintervall für diesen Mittelwert ist [0,78; 0,82].

Tabelle 2: Pearson-Autokorrelation zum Lag  $\tau$  der betrachteten Banken

Bank	Empirische Autokorrelation zum Lag $\tau$									
	$\tau = 1$	2	3	4	5	6	7	8	9	10
BayernLB	0,98	0,96	0,95	0,94	0,93	0,91	0,90	0,89	0,88	0,86
Commerzbank	0,96	0,91	0,88	0,84	0,81	0,77	0,75	0,74	0,72	0,68
Deutsche Bank	0,97	0,94	0,91	0,88	0,85	0,82	0,81	0,79	0,78	0,77
Dresdner Bank	0,98	0,95	0,93	0,90	0,89	0,87	0,86	0,85	0,84	0,82
DZ Bank	0,99	0,98	0,97	0,97	0,96	0,95	0,95	0,94	0,93	0,93
HeLaBa	0,98	0,98	0,97	0,96	0,95	0,95	0,94	0,93	0,92	0,91
HSH Nordbank	0,98	0,98	0,98	0,97	0,97	0,96	0,96	0,95	0,95	0,95
HypoVereinsbank	0,97	0,95	0,93	0,91	0,89	0,87	0,85	0,84	0,84	0,83
LB Berlin	0,98	0,97	0,96	0,95	0,94	0,94	0,94	0,93	0,93	0,92
LBBW	0,98	0,97	0,95	0,94	0,93	0,91	0,90	0,88	0,87	0,85
NordLB	0,97	0,96	0,95	0,94	0,92	0,91	0,91	0,90	0,89	0,88
Postbank	0,98	0,98	0,97	0,96	0,96	0,95	0,95	0,95	0,95	0,95
WestLB	0,98	0,96	0,94	0,91	0,89	0,86	0,84	0,82	0,79	0,77

Tabelle 3: Pearson-Korrelationsmatrix der betrachteten Banken.

BayernLB	1,00	0,78	0,82	0,80	0,86	0,91	0,78	0,92	0,90	0,94	0,89	0,72	0,89
Commerzbank	0,78	1,00	0,83	0,91	0,76	0,69	0,52	0,83	0,74	0,80	0,72	0,58	0,81
Deutsche Bank	0,82	0,83	1,00	0,75	0,90	0,71	0,85	0,95	0,81	0,88	0,87	0,71	0,71
Dresdner Bank	0,80	0,91	0,75	1,00	0,78	0,77	0,52	0,79	0,83	0,76	0,70	0,70	0,86
DZ Bank	0,86	0,76	0,90	0,78	1,00	0,79	0,84	0,89	0,90	0,89	0,92	0,86	0,69
HeLaBa	0,91	0,69	0,71	0,77	0,79	1,00	0,72	0,84	0,89	0,88	0,83	0,74	0,86
HSH Nordbank	0,78	0,52	0,85	0,52	0,84	0,72	1,00	0,86	0,80	0,82	0,87	0,71	0,60
HypoVereinsbank	0,92	0,83	0,95	0,79	0,89	0,84	0,86	1,00	0,88	0,94	0,92	0,71	0,83
LB Berlin	0,90	0,74	0,81	0,83	0,90	0,89	0,80	0,88	1,00	0,87	0,88	0,80	0,81
LBBW	0,94	0,80	0,88	0,76	0,89	0,88	0,82	0,94	0,87	1,00	0,93	0,74	0,83
NordLB	0,89	0,72	0,87	0,70	0,92	0,83	0,87	0,92	0,88	0,93	1,00	0,77	0,71
Postbank	0,72	0,58	0,71	0,70	0,86	0,74	0,71	0,71	0,80	0,74	0,77	1,00	0,55
WestLB	0,89	0,81	0,71	0,86	0,69	0,86	0,60	0,83	0,81	0,83	0,71	0,55	1,00

## 4.2 Auswertung

Wir werden uns hier mit der Anwendung der beschriebenen Methode auf die Daten beschäftigen. Bei der REML-Methode untersuchen wir bei jedem Modell zwei Fälle, die sich mit der Struktur der Reststörgröße (mit und ohne  $AR(1)$ -Fehler) unterscheiden. Diese beide Strukturen lassen sich einfach in SAS mit *Proc Mixed* eingeben (vergleiche Littell u.a. (2007)). Bei dem Bayesianischen Ansatz unterscheiden wir dagegen zwischen einer Normalverteilung und einer Polya-Baum-Verteilung für die zufälligen Effekte. Die PBM-verteilten zufälligen Effekte lassen sich implementieren durch die R-Funktion (*PTlmm*) im R-Paket *DPpackage* (vergleiche Jara u.a. (2011)). Der Gibbs-Sampling-Algorithmus – für normalverteilte zufällige Effekte – lässt sich in *BUGS* (vergleiche Spiegelhalter u.a. (1994-2003)) implementieren. In allen Berechnungen werden wir bei dem Bayesianischen Ansatz 310.000 Iterationen mit einer Aufwärmphase von 10.000 und einen Thinning-Faktor von 30 nutzen. Es ergibt sich eine Markov-Kette vom Umfang 10.000. Die Konvergenz dieser Kette erfolgt durch die Gelman-Rubin-Diagnostik (vergleiche Brooks und Gelman (1998)). Basierend auf dem Output-Objekt aus der Funktion *PTlmm* ist die Konvergenzdiagnostik via das R-Paket *coda* (vergleiche Plummer u.a. (2006)) durchgeführt worden. Die daraus resultierende Potenzial-Reduktion-Faktoren liegen alle bei etwa 1 ( $\hat{R} \approx 1$ ), so dass wir davon ausgehen können, dass die generierten Ketten konvergieren. Für die wichtigsten Modellparameter der Untersuchung werden wir Konfidenzintervalle sowie Intervalle höchster Dichte (highest probability density; HPD) geben. Die HPD-Intervalle lassen sich, wie bei Chen und Shao (1999) beschrieben und im R-Paket *boa* (vergleiche Smith (2007)) implementiert ist, bestimmen. Unser Ziel ist hier, ein optimales Modell zu finden, welches die Längsschnitt- sowie die Querschnittkorrelation realistischerweise modelliert.

Werden nun die Modelle mit festem Unternehmenseffekt angepasst und Schätzer mit der KQ-Methode berechnet, ergibt sich die Tabelle 4. Die Schätzung liefert eine sehr niedrige Verzerrung ( $\text{plim}(\hat{\eta}_{KQ} - \eta) = -0,005$ ), so dass wir davon ausgehen können, dass die Schätzer der dynamischen Modellparameter konsistent sind. In Tabelle 4 wird das niedrigere AIC beziehungsweise größte  $R_{\text{adj}}^2$  beim Modell mit der exogenen Variable *iTraxx* angenommen, sodass dieses Modell vorrangig behandelt wird. Das Modell verbessert zwar die Struktur der Längsschnittkorrelation (Tabelle 8) in entsprechendem statischen Modell, aber seine Schwäche liegen überwiegend in der Querschnittkorrelation. Wir werden für die nächsten Seiten die Modelle mit zufälligen Effekten an die Daten anpassen, in der Hoffnung bessere Ergebnisse zu erzielen. Wir betrachten zunächst die im Abschnitt 3.2.2 beschriebenen Modelle mit zufälligem Unternehmenseffekt.

#### 4 Material und Auswertung

Tabelle 4: Parameterschätzungen bei den Modellen mit festem Unternehmenseffekt.

Modell: $CDS_{it} = \alpha_i^* + \beta iTraxx_t + \eta CDS_{i,t-1} + \varepsilon_{it}$							
	Statisches Modell		Dynamische Modelle				
	$\hat{\alpha}^*$	$\hat{\beta}$	ohne iTraxx		mit iTraxx		
Banken	$\hat{\alpha}^*$	$\hat{\beta}$	$\hat{\alpha}^*$	$\hat{\eta}$	$\hat{\alpha}^*$	$\hat{\beta}$	$\hat{\eta}$
BayernLB	15,13	0,65	1,51	0,98	0,57	0,04	0,95
Commerzbank	10,34	0,65	1,40	0,98	0,30	0,04	0,95
Deutsche Bank	16,22	0,65	1,54	0,98	0,63	0,04	0,95
Dresdner Bank	11,05	0,65	1,30	0,98	0,22	0,04	0,95
DZ Bank	20,03	0,65	1,60	0,98	0,83	0,04	0,95
HeLaBa	20,23	0,65	1,61	0,98	0,85	0,04	0,95
HSH Nordbank	82,89	0,65	2,95	0,98	4,28	0,04	0,95
HypoVereinsbank	7,60	0,65	1,36	0,98	0,17	0,04	0,95
LB Berlin	15,25	0,65	1,53	0,98	0,60	0,04	0,95
LBBW	13,30	0,65	1,48	0,98	0,48	0,04	0,95
NordLB	8,93	0,65	1,38	0,98	0,23	0,04	0,95
Postbank	-1,51	0,65	1,21	0,98	0,28	0,04	0,95
WestLB	31,12	0,65	1,81	0,98	1,41	0,04	0,95
Varianz:							
$\hat{\sigma}_\varepsilon^2$		605,65		44,58			43,29
Bestimmtheitsmaße:							
AIC		47364,00		33915			33776
$R_{adj}^2$		0,92		0,99			0,99



Tabelle 5: Parameterschätzungen und 95% Konfidenzintervalle bez. HPD-Intervalle für Varianzparameter (in Klammern) bei den Modellen mit zufälligem Unternehmenseffekt.

	Modell: $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \alpha_i + \varepsilon_{it}$			
	REML	Normalverteilung	PBM	REML (AR(1)-Fehler)
<b>Statisches Modell</b>				
$\hat{\sigma}_\alpha^2$	424,55 ([217,87; 1162,13])	449,06 ([217,08; 1149,22])	421,13 ([149,14; 786,40])	242,52 ([79,08; 3106,28])
$\hat{\sigma}_\varepsilon^2$	604,65 ([581,84; 628,84])	604,81 ([582,11; 628,04])	605,04 ([581,44; 628,98])	1004 ([722,79; 1491,03])
$\hat{\rho}_{ij}(1, 1)$	0,00	0,00	0,00	0,00
$\hat{\rho}_{ii}(2, 1)$	0,41	0,41 ([0,26; 0,65])	0,35 ([0,27; 0,68])	0,98
$\hat{\mu}$	19,53	19,54	17,26	37,77
$\hat{\beta}$	0,66	0,65	0,65	0,44
$\hat{\phi}$				0,98
AIC	47297,8		47236,12	33262,5
<b>Dynamisches Modell (ohne iTraxx)</b>				
$\hat{\sigma}_\alpha^2$	0,01 ([0,00; 5,03])	0,08 ([0,00; 0,32])	0,78 ([0,15; 1,57])	0,00 ([0,00; 0,00])
$\hat{\sigma}_\varepsilon^2$	44,60 ([42,91; 46,38])	44,58 ([42,90; 46,30])	44,56 ([42,89; 46,33])	44,61 ([42,92; 46,4])
$\hat{\rho}_{ij}(1, 1)$	0,00	0,00	0,00	0,00
$\hat{\rho}_{ii}(2, 1)$	0,70	0,70 ([0,70; 0,70])	0,70 ([0,70; 0,70])	0,68
$\hat{\mu}$	1,27	1,35	1,5	1,16
$\hat{\eta}$	0,98	0,98	0,98	0,98
$\hat{\phi}$				-0,05
AIC	333918,3		47236,12	33906,0
<b>Dynamisches Modell (mit iTraxx)</b>				
$\hat{\sigma}_\alpha^2$	1,02 ([0,47; 3,51])	1,24 ([0,42; 3,06])	1,52 ([0,57; 3,48])	0,90 ([0,41; 3,15])
$\hat{\sigma}_\varepsilon^2$	43,37 ([42,91; 46,38])	43,40 ([41,73; 45,12])	43,38 ([41,74; 45,10])	43,38 ([41,74; 45,12])
$\hat{\rho}_{ij}(1, 1)$	0,00	0,00	0,00	0,00
$\hat{\rho}_{ii}(2, 1)$	0,69	0,69 ([0,69; 0,69])	0,69 ([0,69; 0,70])	0,68
$\hat{\mu}$	0,73	0,73	0,65	0,69
$\hat{\beta}$	0,03	0,03	0,04	0,03
$\hat{\eta}$	0,95	0,95	0,95	0,95
$\hat{\phi}$				-0,03
AIC	333811,1			33806,2

#### 4 Material und Auswertung

Grundsätzlich kann davon ausgegangen werden, dass die aus diesem Abschnitt resultierenden Querschnittskorrelationen Null betragen, da die dafür zuständigen Faktoren ( $\lambda_t, t = 1, \dots, T$ ) nicht im Modell aufgenommen wurden. Werden dort definierte Modelle auf die Daten angewendet, so ergibt sich die Tabelle 5.

Bei den Modellen ohne AR(1)-Fehler kommen drei Algorithmen zur Schätzung der Modellparameter in Frage: der NR-Algorithmus (für den REML-Ansatz), der Gibbs-Sampling-Algorithmus (für normalverteilte zufällige Unternehmenseffekte) sowie ein kombinierter Gibbs-Sampling- und Metropolis-Hasting-Algorithmus (für die Annahme einer PBM-Verteilung an die zufälligen Unternehmenseffekte). Bei den Modellen mit AR(1)-Fehler ist die Analyse – aufgrund der Schwierigkeiten bei der Implementierung der Bayesianischen Methode – auf den REML-Ansatz beschränkt worden. Nichtsdestotrotz dürften sich die REML-Schätzer nicht signifikant von den Bayesianischen Schätzern unterscheiden. Der einzige Nachteil ist, dass Kreditibilitätsintervalle für die entsprechenden Korrelationen nicht bestimmt werden können.

Aus Tabelle 5 – wie auch aus den Tabellen 6 und 7 – ist abzulesen, dass die aus der REML-Methode sowie Bayesianische Methode – unter Annahme einer Normalverteilung – entstehenden Schätzer für die Varianzparameter sehr ähnliche Konfidenzintervalle beziehungsweise HPD-Intervalle liefern. Es gibt zwar (fast) keine signifikanten Unterschiede zwischen den geschätzten Modellparametern, aber die HPD-Intervalle sind unter der Annahme einer PBM-Verteilung (für einige Fälle) nicht so ähnlich wie bei den anderen. Dies könnte daran liegen, dass die Verteilung der zufälligen Effekten von der Normalverteilung abweicht, und würde die bereits erwähnte Flexibilität der PBM-Verteilung gegenüber der Normalverteilung (siehe Abschnitt 2.2) bestätigen. Außerdem ist auch zu erkennen, dass die Modelle mit AR(1)-Fehler die entsprechenden Modelle ohne AR(1)-Fehler im Sinn des AIC stets verbessern, was der Längsschnittkorrelation (Tabelle 8 und 9) ermöglicht, sich der empirischen Autokorrelation (Tabelle 2) zu nähern, aber einer Inflation der Fehlervarianz – insbesondere bei den statischen Modellen – zufolge hat, die bewirkt, dass die Querschnittskorrelation Null wird.

Von besonderem Interesse bei der Untersuchung ist die Schätzung von  $\sigma_\lambda^2$ . Wie aus dem Abschnitt 3 ersichtlich wird, ist  $\sigma_\lambda^2$  für die Korrelation von Ausfallwahrscheinlichkeiten zwischen Unternehmen verantwortlich. Tabellen 6 und 7 zeigen zunächst, dass ohne Verwendung von verzögerten CDS-Preisen aufgrund der erhöhten Fehlervarianz die Querschnittskorrelation unter 13% liegt. Diese Werte verdreifachen sich bei der Aufnahme von verzögerten CDS-Preisen im Modell.

Tabelle 6: Parameterschätzungen und 95% Konfidenzintervalle bez. HPD-Intervalle für Varianzparameter (in Klammern) bei den Modellen mit zufälligem Zeiteffekt.

	Modell: $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \lambda_t + \varepsilon_{it}$			
	REML	Normalverteilung	PBM	REML(AR(1)-Fehler)
<b>Statisches Modell</b>				
$\hat{\sigma}_\lambda^2$	91,44 ([72,36; 119,27])	92,34 ([70,73; 116,86])	90,62 ([90,62; 90,62])	5,86 ([4,83; 7,25])
$\hat{\sigma}_\varepsilon^2$	905,70 ([870,21; 943,37])	906,04 ([869,62; 942,65])	866,57 ([833,07; 899,78])	1161,9 ([852,33; 1678,02])
$\hat{\rho}_{ij}(1, 1)$	0,09	0,09 ([0,07; 0,11])	0,09 ([0,08; 0,10])	$\approx 0$
$\hat{\rho}_{ii}(2, 1)$	0,00	0,00	0,00	0,98
$\hat{\mu}$	19,54	19,55	19,52	26,61
$\hat{\beta}$	0,66	0,65	0,65	0,51
$\hat{\phi}$				0,98
AIC	49629,4			32431,3
<b>Dynamisches Modell (ohne iTraxx)</b>				
$\hat{\sigma}_\lambda^2$	16,00 ([13,73; 18,90])	16,12 ([16,06; 16,18])	15,95 ([15,95; 15,95])	16,02 ([13,74; 18,91])
$\hat{\sigma}_\varepsilon^2$	28,65 ([27,53; 29,84])	28,64 ([27,51; 29,84])	26,71 ([25,72; 27,80])	28,65 ([27,52; 29,86])
$\hat{\rho}_{ij}(1, 1)$	0,36	0,36 ([0,34; 0,37])	0,37 ([0,36; 0,38])	0,36
$\hat{\rho}_{ii}(2, 1)$	0,70	0,70 ([0,70; 0,70])	0,70 ([0,70; 0,70])	0,68
$\hat{\mu}$	1,32	1,32	1,26	1,13
$\hat{\eta}$	0,98	0,98	0,98	0,98
$\hat{\phi}$				-0,09
AIC	32483,7			32447,6
<b>Dynamisches Modell (mit iTraxx)</b>				
$\hat{\sigma}_\lambda^2$	15,58 ([13,36; 18,42])	15,65([13,30; 18,30])	15,58 ([13,36; 18,42])	15,68([13,45; 18,53])
$\hat{\sigma}_\varepsilon^2$	28,65 ([27,53; 29,85])	28,66([27,51; 29,84])	28,59 ([27,42; 29,74])	28,65([27,52; 29,85])
$\hat{\rho}_{ij}(1, 1)$	0,36	0,36 ([0,31; 0,39])	0,35 ([0,30; 0,39])	0,36
$\hat{\rho}_{ii}(2, 1)$	0,70	0,70([0,70; 0,70])	0,67 ([0,65; 0,71])	0,68
$\hat{\mu}$	0,12	0,11	0,27	0,06
$\hat{\beta}$	0,01	0,01	0,01	0,01
$\hat{\eta}$	0,98	0,98	0,98	0,98
$\hat{\phi}$				-0,08
AIC	32483,1			32449,1

#### 4 Material und Auswertung

Tabelle 7: Parameterschätzungen und 95% Konfidenzintervalle bez. HPD-Intervalle für Varianzparameter (in Klammern) bei den Modellen mit zufälligen Unternehmens- und Zeiteffekten.

Modell: $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{it}$			
	REML	Normalverteilung	REML(AR(1)-Fehler)
<b>Statisches Modell</b>			
$\hat{\sigma}_\alpha^2$	424,86 ([218,12; 1161,91])	506,41 ([216,82; 1156,64])	232,42 ([71,59; 4030,57])
$\hat{\sigma}_\lambda^2$	124,13 ([104,27; 150,28])	124,99 ([103,77; 149,05])	5,87 ([4,84; 7,28])
$\hat{\sigma}_\varepsilon^2$	480,82 ([461,96; 500,85])	480,96 ([462,43; 501,08])	921,12 ([639,74; 1440,43])
$\hat{\rho}_{ij}(1, 1)$	0,12	0,11 ([0,06; 0,16])	$\approx 0$
$\hat{\rho}_{ii}(2, 1)$	0,41	0,41 ([0,26; 0,65])	0,98
$\hat{\mu}$	19,54	19,49	29,68
$\hat{\beta}$	0,66	0,65	0,51
$\hat{\phi}$			0,98
<i>AIC</i>	46707,5		32431,5
<b>Dynamisches Modell (ohne iTraxx)</b>			
$\hat{\sigma}_\alpha^2$	0,20 ([0,07; 1,38])	0,25 ([0,03; 0,75])	0,14 ([0,05; 1,24])
$\hat{\sigma}_\lambda^2$	16,11 ([13,82; 19,02])	16,19 ([13,75; 18,96])	16,07 ([13,79; 18,97])
$\hat{\sigma}_\varepsilon^2$	28,49 ([27,37; 29,68])	28,51 ([27,42; 29,67])	28,52 ([27,39; 29,72])
$\hat{\rho}_{ij}(1, 1)$	0,36	0,36 ([0,32; 0,40])	0,36
$\hat{\rho}_{ii}(2, 1)$	0,70	0,70 ([0,70; 0,71])	0,69
$\hat{\mu}$	1,93	1,93	1,65
$\hat{\eta}$	0,98	0,98	0,98
$\hat{\phi}$			-0,09
<i>AIC</i>	32478,0		32443,7
<b>Dynamisches Modell (mit iTraxx)</b>			
$\hat{\sigma}_\alpha^2$	0,39 ([0,17; 1,74])	0,49 ([0,13; 1,30])	0,28 ([0,11; 1,37])
$\hat{\sigma}_\lambda^2$	15,23 ([13,05; 18,01])	15,32 ([12,97; 18,05])	15,37 ([13,17; 18,18])
$\hat{\sigma}_\varepsilon^2$	28,46 ([27,34; 29,64])	28,48 ([27,34; 29,66])	28,46 ([27,34; 29,66])
$\hat{\rho}_{ij}(1, 1)$	0,35	0,35 ([0,30; 0,38])	0,35
$\hat{\rho}_{ii}(2, 1)$	0,70	0,70 ([0,69; 0,71])	0,69
$\hat{\mu}$	0,36	0,37	0,27
$\hat{\beta}$	0,02	0,02	0,02
$\hat{\eta}$	0,97	0,97	0,97
$\hat{\phi}$			-0,08
<i>AIC</i>	32467,0		32437,4

#### 4 Material und Auswertung

Auffällig bei der Analyse ist, dass die unternehmensspezifischen zufälligen Effekte bei dynamischen Modellen auf etwa Null geschätzt sind. Dies deutet darauf hin, dass die verzögerte endogene Variable die Information über die möglichen unternehmensspezifischen – erklärenden – Variablen enthält, was impliziert, dass zeitspezifische Variablen wahrscheinlich die Einzigen sind, die bei der Untersuchung vernachlässigt wurden. Dies steht im Zusammenhang mit Norden und Wagner (2008), die bereits angaben, dass die wesentlichen und relevanten Informationen für das Kreditgeschäft in den CDS-Preisen enthalten sind. Außerdem ist es von Interesse zu betonen, dass die Ausfallwahrscheinlichkeitskorrelation stark mit der Modellwahl verbunden ist. In der Analyse haben statischen Modelle mit zufälligem Zeiteffekt sowie zufälligen Unternehmens- und Zeiteffekten jeweils mit AR(1)-Fehler das niedrigste AIC erreicht. Somit könnten sie als beste Modelle angesehen werden. Allerdings beträgt die Querschnittkorrelation bei diesen beiden Modellen etwa Null, während die gemittelte paarweise empirische Korrelation zwischen betrachteten Unternehmen (Tabelle 2) 0,8 beträgt.

Eine sinnvolle Modellwahl würde erstens auf dem dynamischen Modell mit zufälligen zeitspezifischen Effekten ( $CDS_{it} = \mu + \eta CDS_{i,t-1} + \lambda_t + \varepsilon_{it}$ ) und zweitens auf dem dynamischen Modell mit zufälligem Zeiteffekt und dem Marktindex (iTraxx) als Kovariable ( $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \lambda_t + \varepsilon_{it}$ ) aus den folgenden Gründen fallen:

- Diese Modelle weisen eine perfekte Anpassung an die empirisch beobachteten Längsschnittkorrelation (genau wie die Besten) auf.
- Bei den genannten statischen Modellen können zwei Unternehmen – theoretisch – nur am gleichen Tag korreliert werden, obwohl die Längsschnittkorrelation – auch wenn der Abstand zwischen Handelstagen 10 ist – nicht Null ist. Die vorgeschlagenen Modelle lösen dieses Problem indem die Querschnittkorrelation sich theoretisch nach der Faustregel  $\hat{\rho}_{ij}(t, s) = \hat{\rho}_{ii}(t, s)\hat{\rho}_{ij}(t, t)$  richtet (siehe Formel (87) sowie Tabelle 9). Somit sinken die Querschnitt- als auch die Längsschnittkorrelationen mit wachsendem Abstand zwischen den Handelstagen.
- Es ergibt sich bei diesen Modellen ein plausibles Zusammenwirken der Längs- und Querschnittkorrelation. Die Querschnittkorrelation ist zwar kleiner als die Empirische, ist aber besser im Vergleich zu der aus diesen statischen Modellen entstehenden Korrelation.

Diese beiden Modelle erzielen die gleichen Schätzer für die Korrelation und unterscheiden sich nur noch von der Parameteranzahl. Interessant bei der ersten Wahl

#### *4 Material und Auswertung*

(neben der kleinsten Parameteranzahl) ist auch hier, dass keine andere Variable benötigt wird, um die CDS-Prämie betrachteter Banken erklären zu können.

Die Korrelationseffekte lassen sich als mögliche nicht diversifizierbarer Risikoanteil identifizieren, die die CDS-Preise aller Unternehmen beeinflussen. Das Ergebnis aus Tabellen 8 und 9 stellt klar und deutlich dar, dass CDS-Preise weitgehend von unsystematischen, unternehmensspezifischen Faktoren getrieben werden, da die Längsschnittkorrelation besonders höher als die Querschnittkorrelation ist. Dies steht vermutlich im Zusammenhang mit der Erfahrung bei Schieflagen und Insolvenzen vieler Unternehmen, bei denen Managementfehler und Unregelmäßigkeiten als Hauptursachen identifiziert werden konnten (vergleiche Hamerle und Rösch (2003)).

#### 4 Material und Auswertung

Tabelle 8: Korrelationsschätzungen bei den Modellen mit festem und zufälligem Unternehmenseffekt

$s$	1	1	1	1	1	1	1	1	1	1
$t$	1	2	3	4	5	6	7	8	9	10
<b>Modell: <math>CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \varepsilon_{it}</math></b>										
<b>Statisches Modell</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Dynamisches Modell</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,69	0,55	0,46	0,40	0,36	0,32	0,29	0,27	0,25
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Modell: <math>CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \alpha_i + \varepsilon_{it}</math></b>										
<b>Statisches Modell</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,41	0,41	0,41	0,41	0,41	0,41	0,41	0,41	0,41
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Statisches Modell mit AR(1)-Fehler</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,98	0,96	0,94	0,92	0,90	0,88	0,86	0,85	0,83
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Dynamisches Modell</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,70	0,56	0,48	0,42	0,38	0,35	0,32	0,30	0,28
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
<b>Dynamisches Modell mit AR(1)-Fehler</b>										
$\hat{\rho}_{ii}(t, s)$	1,00	0,84	0,79	0,76	0,74	0,73	0,72	0,71	0,71	0,70
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

4 Material und Auswertung

Tabelle 9: Korrelationsschätzungen bei den Modellen mit zufälligem Zeiteffekt sowie zufälligen Unternehmens- und Zeiteffekten

s	1	1	1	1	1	1	1	1	1	1
t	1	2	3	4	5	6	7	8	9	10
Modell: $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \lambda_t + \varepsilon_{it}$										
Statisches Modell										
$\hat{\rho}_{ii}(t, s)$	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,09	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Statisches Modell mit AR(1)-Fehler										
$\hat{\rho}_{ii}(t, s)$	1,00	0,98	0,96	0,94	0,92	0,90	0,89	0,87	0,85	0,83
$\hat{\rho}_{ij, i \neq j}(t, s)$	$\approx 0$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Dynamisches Modell										
$\hat{\rho}_{ii}(t, s)$	1,00	0,70	0,55	0,46	0,40	0,36	0,32	0,29	0,27	0,25
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,36	0,25	0,20	0,17	0,15	0,14	0,13	0,12	0,11	0,10
Dynamisches Modell mit AR(1)-Fehler										
$\hat{\rho}_{ii}(t, s)$	1,00	0,68	0,56	0,49	0,44	0,40	0,37	0,35	0,33	0,31
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,36	0,26	0,22	0,19	0,17	0,16	0,15	0,14	0,14	0,13
Modell: $CDS_{it} = \mu + \beta iTraxx_t + \eta CDS_{i,t-1} + \alpha_i + \lambda_t + \varepsilon_{it}$										
Statisches Modell										
$\hat{\rho}_{ii}(t, s)$	1,00	0,41	0,41	0,41	0,41	0,41	0,41	0,41	0,41	0,41
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Statisches Modell mit AR(1)-Fehler										
$\hat{\rho}_{ii}(t, s)$	1,00	0,98	0,96	0,94	0,92	0,90	0,89	0,87	0,85	0,84
$\hat{\rho}_{ij, i \neq j}(t, s)$	$\approx 0$	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Dynamisches Modell										
$\hat{\rho}_{ii}(t, s)$	1,00	0,70	0,57	0,48	0,43	0,39	0,35	0,33	0,31	0,29
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,35	0,24	0,19	0,16	0,14	0,13	0,12	0,11	0,10	0,09
Dynamisches Modell mit AR(1)-Fehler										
$\hat{\rho}_{ii}(t, s)$	1,00	0,69	0,57	0,49	0,44	0,40	0,37	0,35	0,33	0,31
$\hat{\rho}_{ij, i \neq j}(t, s)$	0,35	0,25	0,21	0,19	0,17	0,16	0,15	0,14	0,13	0,13



## 5 Diskussion

Die Schätzung von Ausfallwahrscheinlichkeitskorrelationen erfolgt in dieser Arbeit durch balancierte Panel-Regressionen von CDS-Preisen. Unterschiedliche Annahmen an die Störgrößen sind gemacht worden und zwei Schätzverfahren darauf angewendet worden. Dies sind die REML-Methode sowie die Bayesesianische Methode. Bei dem REML-Ansatz wurde die Annahme über die Normalverteilung an die Störgröße getroffen. In diesem Fall wurde die Existenz und die statistischen Eigenschaften der REML-Schätzer bei  $AR(1)$ -Fehlern untersucht. Bei den bayesianischen Ansatz sind die Normalverteilung sowie PBM-Verteilung in Frage gekommen. Allerdings ist der Fall der  $AR(1)$ -Fehler aufgrund von Schwierigkeiten bei der Implementierung nicht berücksichtigt worden. Werden für die Fehler eine Struktur eines  $AR(1)$ -Prozesses sowie nicht informative a-priori-Verteilungen für die Modellparameter – bei normalverteilten zufälligen Effekten – angenommen, so ist das Problem der Zulässigkeit der a-posteriori-Verteilung noch offen. Man könnte dafür die Integrierbarkeit von  $|X'\Sigma^{-1}X|^{-\frac{1}{2}}|\Sigma|^{-\frac{1}{2}}$  über  $\phi$  untersuchen, da die Abschätzung  $\int_{-1}^1 \int_{\mathbb{R}^{(\sum_{i=1}^p n_i)}} \int_{\mathbb{R}^m} f(y|\delta, \xi, \sigma_0^2) f(\xi|\sigma_1^2, \dots, \sigma_p^2) d\delta d\xi d\phi \leq \int_{-1}^1 \frac{|X'\Sigma^{-1}X|^{-\frac{1}{2}}|\Sigma|^{-\frac{1}{2}}}{(2\pi\sigma_0^2)^{\frac{N \cdot T - m}{2}}} d\phi$  aus Satz 1 von Sun u.a. (2001) hergeleitet werden kann.

Das beste Modell – im Sinn des AIC – für diese Analyse führt zwar zu einer annähernd perfekten Modellierung der Längsschnittkorrelation, schätzt aber die Querschnittkorrelation auf etwa Null, obwohl andere betrachtete Modelle diese Korrelation auf 0,36 schätzen. Dies stellt die Frage nach der Plausibilität des AIC als Kriterium für den Modellvergleich bei Panel-Regressionen im Rahmen der Analyse der Korrelation. Die dynamischen Modelle weisen mit zufälligem Zeiteffekt 0,70 und 0,36 die beste Anpassung an beide Korrelationen auf. Außerdem ergibt sich ausschließlich bei diesen Modellen auch theoretisch ein plausibles Zusammenwirken der Längs- und Querschnittkorrelation.

Noch offen ist der Grund, warum sich die aus den betrachteten Modellen resultierenden Querschnittkorrelationen signifikant von der empirischen unterscheiden.

Collin-Dufresne (2009) erklärte zum Beispiel, dass während der Subprime-Krise Copula-Modelle bei  $LGD = 60\%$  zu einer unzulässigen Korrelation (das heißt, dass die Korrelation nicht im Intervall  $[-1; 1]$  lag) führte, sodass man  $LGD = 1$  annehmen musste, um zulässige Korrelationen zu erhalten. Selbst bei dieser Annahme blieb das Modell instabil. Dies ist bedenklich, da selbst zulässige Korrelationen, die unbestritten bei unseren Modellen – und zwar mit den Daten mit demselben Hintergrund – sind, eine Herausforderung bei dem Copula-Modell darstellten. Das von uns vorgeschlagene Modell könnte daher als ein sehr gute Alternative angesehen

## 5 *Diskussion*

werden, da eine zulässige und vernünftige Korrelation ohne unternehmensspezifische Variablen erreicht wird, die als wichtige Inputparameter für das Copula-Modell identifiziert worden sind. Um eine genaue Aussage über die Güte beider Modelle zu machen, werden wir in unserer künftigen Forschung das Copula-Modell auch auf unsere Daten anwenden.

Ratingmigrationshistorien könnten – trotz ihrer im Kapitel 3 genannten Mängeln – eine Alternative für die in dieser Studie betrachteten Messungen (CDS-Preise) zur Quantifizierung von Ausfallwahrscheinlichkeiten sein (vergleiche Kim u.a. (2012)). Inwieweit einer der beiden Messungen adäquater für die Ausfallkorrelationen ist, werden wir noch künftig untersuchen.

## Tabellenverzeichnis

1	Bilanzsummen der betrachteten Banken. . . . .	74
2	Empirische Autokorrelation. . . . .	81
3	Empirische Korrelation. . . . .	82
4	Modelle mit festem Unternehmenseffekt. . . . .	84
5	Modell mit zufälligem Unternehmenseffekt. . . . .	85
6	Modell mit zufälligem Zeiteffekt. . . . .	87
7	Modell mit zufälligen Unternehmens- und Zeiteffekt. . . . .	88
8	Korrelationschätzungen bei den Modellen mit festem und zufälligem Unternehmenseffekt . . . . .	91
9	Korrelationschätzungen bei den Modellen mit zufälligem Zeiteffekt sowie zufälligen Unternehmens- und Zeiteffekten . . . . .	92

## **Abbildungsverzeichnis**

1	Zahlungsströme eines CDS ohne Kreditausfall . . . . .	3
2	Zahlungsströme eines CDS bei Kreditausfall . . . . .	3
3	Zeitreihen der CDS-Preise . . . . .	76
4	Zeitreihen der CDS-Preise und Markit-iTraxx-Europe-Indices . . . .	77
5	Zeitreihen der der prozentualen Änderung in Fünf-Jahres-CDS-Preisen	79

## **Abkürzungsverzeichnis**

<b>AG</b>	Aktiengesellschaft
<b>AIC</b>	<i>Akaike Information Criterion</i>
<b>AöR</b>	Anstalt des öffentlichen Rechts
<b>B</b>	Anleihe
<b>BayernLB</b>	Bayerische Landesbank AöR
<b>Basel II</b>	Basler Ausschusses für Bankenaufsicht II
<b>CDO</b>	Collateralized Debt Obligation
<b>CDS</b>	<i>Credit Default Swap</i>
<b>CreditVaR</b>	Credit Value at Risk
<b>Deka Bank</b>	Deka Bank Deutsche Girozentrale AöR
<b>dim</b>	Dimension
<b>DPM</b>	Dirichlet-Prozess-Mischung
<b>E</b>	Erwartungswert
<b>EM</b>	Expectation-Maximization
<b>FS</b>	Fischer Scoring
<b>HeLaBa</b>	Landesbank Hessen-Thüringen Girozentrale AöR
<b>HPD</b>	Highest probability density
<b>HRE</b>	Hypo Real Estate Holding AG
<b>HypoVereinsbank</b>	Bayerische Hypo- und Vereinsbank AG
<b>I</b>	Einheitsmatrix
<b>iTraxx</b>	CDS-Index in den Regionen Europa Asien
<b>IG</b>	Inverse Gamma

## Abkürzungsverzeichnis

<b>KfW</b>	Kreditanstalt für Wiederaufbau AöR
<b>kov</b>	Kovarianzfunktion
<b>korr</b>	Korrelationsfunktion
<b>KQ</b>	Kleinste Quadrate
<b>L</b>	Verlust
<b>LB Berlin</b>	Landesbank Berlin Holding AG
<b>LBBW</b>	Landesbank Baden-Württemberg AöR
<b>LGD</b>	Verlustrisikoprämie, <i>Loss Given Default</i>
$l_{R,p}$	Restringierte-Profil-Maximum-Likelihood
<b>MBP</b>	Maximization by Parts
<b>MINQU</b>	Minimum Norm Quadratic Unbiased
<b>ML</b>	Maximum Likelihood
<b>MPBM</b>	multivariate Polya-Baum-Mischung
<b>NordLB</b>	Norddeutsche Landesbank Girozentrale AöR
<b>NR</b>	Newton-Raphson
<b>N</b>	Normalverteilung
<b>PBM</b>	Polya-Baum-Mischung
<b>PD</b>	Ausfallwahrscheinlichkeit, <i>Probability of Default</i>
$R_{adj}^2$	Adjustiertes Bestimmtheitsmaß
<b>REML</b>	Restringierte Maximum Likelihood
<b>sp</b>	CDS-Preise, <i>spread</i>
<b>spur</b>	Spur
<b>SoFFin</b>	Finanzmarktstabilisierungsfonds, auch: Sonderfonds Finanzmarktstabilisierung

*Abkürzungsverzeichnis*

**var**

Varianz

**WGZ**

Westdeutsche Genossenschafts-Zentralbank AG

## Literatur

- T. W. Anderson. Estimation of covariance matrices which are linear combinations or whose inverses are linear combinations of given matrices. *Essays in Probability and Statistics*, pages 1–24, 1970. Statistical Publishing Society, Calcutta, 1969, and University of North Carolina Press, Chapel Hill.
- M. Arellano. *Panel Data Econometrics*. Oxford University Press, 2003.
- S. F. Arnold. *The Theory of Linear Models and Multivariate Analysis*. Wiley, 1981.
- B. H. Baltagi. *Econometric Analysis of Panel Data*. Wiley, 2003.
- C. Bassi Mbobda, C. Lisakowski, F. Müller, N. Reckmann und A. Sommer. Steinbrück Mitgift. Beitrag zum Postbank Finance Award 2009, Technische Universität Dortmund, 2009. URL <http://faz.net/>.
- C. A. Bassi Mbobda. Korrelation von Ausfallwahrscheinlichkeiten. Masterarbeit, TU Dortmund, 2010. Unveröffentlicht.
- D. Birkes und S. S. Wulff. Existence of maximum likelihood estimates in normal variance-components models. *Journal of Statistical Planning and Inference*, 113(1): 35–47, 2003.
- F. Black und J. C. Cox. Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31(2):351–367, 1976.
- C. Bluhm, L. Overbeck und C. Wagner. *An introduction to credit risk modeling*. Chapman & Hall/CRC, Boca Raton, Florida, USA, 2003.
- D. Brigo und K. Chourdakis. Counterparty risk for credit default swap: impact of spread volatility and default correlation. *International Journal of Theoretical and Applied Finance*, 12(7):1007–1026, 2009.
- D. Brigo und F. Mercurio. *Interest Rate Models - Theory and Practice*. New York, springer edition, 2006.
- S. P. Brooks und A. Gelman. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, 1998.
- J. Cariboni, S. Maccaferri und Schoutens W. Lévy processes and the financial crisis: can we design a more effective deposit protection? Working paper, 2011. URL <http://isi2011.congressplanner.eu/pdfs/650291.pdf>.



## Literatur

- S. Chatterjee und A. S. Hadi. *Sensitivity Analysis in Linear Regression*. Wiley, 1988.
- M-H. Chen und Q-M. Shao. Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*, 8(1):69–92, 1999.
- U. Cherubini, S. Mulinacci, F. Gobbi und S. Romagnoli. *Dynamic Copula Methods in Finance*. Wiley Finance Series, 2011.
- P. Collin-Dufresne. A Short Introduction to Correlation Markets. *Journal of Financial Econometrics*, 7(1):12–29, 2009.
- P. Collin-Dufresne, R.S. Goldstein und Martin S. J. The Determinants of Credit Spread Changes. *Journal of Finance*, 56(6):2177–2207, 2001.
- M. K. Cowles und B. P. Carlin. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Journal of the American Statistical Association*, 91(434):883–904, 1996.
- D. Creal, S. J. Koopman und A. Lucas. A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations. *Journal of Business and Economic Statistics*, 29(4):552–563, 2011.
- D. Creal, S. J. Koopman und A. Lucas. Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics*, 2012. Erscheint.
- N. Cressie und N. L. Soumendra. The asymptotic distribution of REML estimator. *Journal of Multivariate Analysis*, 45(2):217–233, 1993.
- V. De Oliveira und M. Ferreira. Maximum likelihood and restricted maximum likelihood estimation for a class of Gaussian Markov random fields. *Metrika*, 74(2):167–183, 2011.
- E. Demidenko und H. Massam. On the existence of the maximum likelihood estimate in variance components models. *Indian Journal of Statistics*, 61(3):431–443, 1999.
- Eugene Demidenko. *Mixed Models Theory and Applications*. Wiley, 2004.
- Die Zeit. HSH Nordbank zahlte unnötig Millionen. BANKENKRISE, 2009. URL <http://www.zeit.de/wirtschaft/unternehmen/2009-09/nordbank-zahlt-millionen>.
- N. R. Draper und H. Smith. *Applied Regression Analysis*. Wiley, 1998.

## Literatur

- M. Durban, J. Harezlak, M.P. Wand und R.J. Carroll. Simple fitting of subject-specific curves for longitudinal data. *Statistic in Medicine*, 24(1):1153–1167, 2005.
- J. Ericsson, Jacobs K. und R. Oviedo. The Determinants of Credit Default Swap Premia. *Journal of Financial and Quantitative Analysis*, 44(1):109–132, 2009.
- L. Fahrmeir und S. Lang. Bayesian semiparametric regression analysis of multi-categorical time-space data. *Annals of the Institute of Statistical Mathematics*, 53(1): 11–30, 2001.
- P. Feldhütter und M. S. Nielsen. Systematic and Idiosyncratic Default Risk in Synthetic Credit Markets. *Journal of Financial Econometrics*, 10(2):292–324, 2012.
- T. S. Ferguson. A Bayesian analysis of some nonparametric problems. *Annals of Statistics*, 1(2):209–230, 1973.
- T. S. Ferguson. Prior Distributions on Spaces of Probability Measures. *Annals of Statistics*, 2(4):615–629, 1974.
- Fitch Ratings. CDS Spreads and Default Risk. Special Report., 2010. URL [http://inwestycje.pl/resources/Attachment/2010/10\\_13/file8529.pdf](http://inwestycje.pl/resources/Attachment/2010/10_13/file8529.pdf).
- K. J. Forbes und R. Rigobon. No Contagion, Only Interdependence: Measuring Stock Market Comovements. *Journal of Finance*, 57(5):2223–2261, 2002.
- Basler Ausschuss für Bankenaufsicht, editor. *Internationale Konvergenz der Eigenkapitalmessung und der Eigenkapitalanforderungen. Überarbeitete Rahmenvereinbarung*. Bank für Internationalen Zahlungsausgleich, Basel, 2004.
- Wayne A. Fuller. *Introduction to Statistical Time Series*. Wiley, 1976.
- A. E. Gelfand, S. E. Hills, A. Racine-Poon und A. F. M. Smith. Illustration of Bayesian Inference in Normal Data Models Using Gibbs Sampling. *Journal of the American Statistical Association*, 85(412):972–985, 1990.
- A. Hamerle und D. Rösch. Risikofaktoren und Korrelationen für Bonitätsveränderungen. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung*, 55(1):199–223, 2003.
- T. E. Hanson. Inference for Mixtures of Finite Polya Tree Models. *Journal of the American Statistical Association*, 101(476):1548–1565, 2006.

## Literatur

- T. E. Hanson und W. Johnson. Bayesian nonparametric multiple imputation of partially observed data with ignorable nonresponse. *Biometrika*, 89(3):529–538, 2002.
- D. A. Harville. Bayesian inference for variance components using only error contrasts. *Biometrika*, 61(2):383–385, 1974.
- F. Heinzl, L. Fahrmeir und T. Kneib. Additive mixed models with Dirichlet process mixture and P-spline priors. *Asta advances in statistical analysis*, 96(1):47–68, 2012.
- J. P. Hobert und G. Casella. The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91(436):1461–1473, 1996.
- C. Hsiao. *Analysis of Panel Data*. Cambridge University Press, 2003.
- C. Hsiao, M. H. Pesaran und A. K. Tahmiscioglu. Bayes estimation of short run coefficients in dynamic panel data Models. Technical document, 1997. URL <http://www.eprg.org.uk/faculty/pesaran/hsiao.pdf>.
- R. M. Huggins. On the robust Analysis of variance components models for pedigree data. *Australian Journal of Statistics*, 35(1):43–57, 1993.
- J. Hull, M. Predescu und A. White. The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking and Finance*, 28(11):2789–2811, 2004.
- A. T. James. Normal multivariate analysis and the orthogonal group. *Annals of Mathematical Statistics*, 25(1):40–75, 1954.
- A. Jara, T. E. Hanson und E. Lesaffre. Robustifying generalized linear mixed models using a new class of mixtures of multivariate Polya trees. *Journal of Computational and Graphical Statistics*, 18(4):838–860, 2009.
- A. Jara, T. E. Hanson, F. A. Quintana, P. Müller und G. L. Rosner. DPpackage: Bayesian Non- and Semi-parametric Modelling in R. *Journal of Computational and Graphical Statistics*, 40(5):1–30, 2011.
- P. Jorion. *Financial Risk Manager Handbook*. Wiley Finance, 6 edition, 2010.
- P. Jorion und G. Zhang. Good and bad creditcontagion: Evidence from credit default swaps. *Journal of Financial Economics*, 84(3):860–883, 2006.

## Literatur

- R. E. Kass und G. Casella. The selection of prior distributions by formal rules. *Journal of American Statistical Association*, 91(435):1343–1370, 1996.
- N. M. Kiefer. Estimation of fixed effect models for time series of cross-sections with arbitrary intertemporal covariance. *Journal of Econometrics*, 14(2):195–202, 1980.
- Y. Kim, L. James und R. Weißbach. Bayesian Analysis of Multi-State Event History Data: Beta-Dirichlet Process Prior. *Biometrika*, 99(1):127–140, 2012.
- H. Kuck. Die 100 größten deutschen Kreditinstitute. *Die Bank. Zeitschrift für Bankpolitik und Bankpraxis*, pages 28–29, 2009.
- N. M. Laird und J. H. Ware. Random-effects models for longitudinal data. *Biometrics*, 38(4):963–974, 1982.
- T. Lancaster. The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2):391–413, 2000.
- Tony Lancaster. *Introduction to Modern Bayesian Econometrics*. Wiley-Blackwell, 2004.
- D. Lando und T. M. Skødeberg. Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking & Finance*, 26(2–3):423–444, 2002.
- M. Lavine. Some aspects of polya tree distributions for statistical modelling. *Annals of Statistics*, 20(3):1222–1235, 1992.
- M. Lavine. Aspects of polya tree distributions for statistical modelling. *Annals of Statistics*, 22(3):1161–1176, 1994.
- D. X. Li. On Default Correlation: A Copula Function Approach. *Journal of Fixed Income*, 9(4):43–54, 2000.
- X. Lin und D. Zhang. Inference in generalized additive mixed models by using smoothing splines. *Journal of the Royal Statistical Society*, 61(2):381–400, 1999.
- M.J. Lindstrom und D.M. Bates. Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data. *Journal of the American Statistical Association*, 83(404):1014–1022, 1988.
- R. C. Littell, G. A. Milliken, W. W. Stroup, R. D. Wolfinger und O. Schabenberger. *SAS for Mixed Model*. SAS, 2007.
- F. A. Longstaff, Mithal S. und E. Neis. Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *Journal of Finance*, 60(5):2213–2253, 2005.

## Literatur

- K. V. Mardia und R. J. Marshall. Maximum Likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, 71(1):135–146, 1984.
- A. J. McNeil, R. Frey und P. Embrechts. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, 2005.
- R. Merton. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, 29(2):449–470, 1974.
- L. Norden und W. Wagner. Credit derivatives and loan pricing. *Journal of Banking and Finance*, 32(12):2560–2569, 2008.
- S. M. Paddock. Bayesian nonparametric multiple imputation of partially observed data with ignorable nonresponse. *Biometrika*, 89(3):529–538, 2002.
- S. M. Paddock, F. Ruggeri und M. Lavine. Randomised polya tree models for nonparametric bayesian inference. *Statistica Sinica*, 13(1):443–460, 2003.
- H. D. Patterson und R. Thompson. Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58(3):545–554, 1971.
- M. Plummer, N. Best, K. Cowles und K. Vines. coda: Convergence Diagnosis and Output Analysis for MCMC. *R News*, 6(1):7–11, 2006. URL <http://CRAN.R-project.org/doc/Rnews/>.
- C. R. Rao und J. Kleffe. *Estimation of Variance Components and Applications*. North Holland, Amsterdam, 1988.
- A. M. Richardson und A. H. Welsch. Asymptotic properties of restricted maximum likelihood (REML) estimates for hierarchical mixed linear models. *Australian Journal of Statistics*, 36(1):31–43, 1994.
- B. Rosenow und R. Weißbach. Modelling correlations in credit portfolio risk. *Journal of Risk Management in Financial Institutions*, 3(1):16–30, 2009.
- D. Ruppert, M.P. Wand und R.J. Carroll. *Semiparametric regression*. Cambridge University Press, 2003.
- S. R. Searle, G. Casella und C. E. McCulloch. *Variance components*. Wiley, 1992.
- B. J. Smith. boa: An R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21(11):1–37, 2007.

## Literatur

- P. X.-K. Song, Y. Fan und Kalbfleisch J. D. Maximization by Parts in Likelihood Inference. *Journal of the American Statistical Association*, 100(472):1145–1158, 2005.
- D. J. Spiegelhalter, A. Thomas, N. G. Best, W. R. Gilks und D. Lunn. Bayesian inference using Gibbs sampling. MRC Biostatistics Unit, Cambridge, England, 1994-2003. URL [www.mrc-bsu.cam.ac.uk/bugs/](http://www.mrc-bsu.cam.ac.uk/bugs/).
- Standard & Poors. Request For Comment: Refinement Of Global CDO Cash Flow Modeling Assumptions. S&P Credit Research, 2006.
- D. Sun, R. K. Tsutakawa und Z. He. Propriety of posteriors with improper priors in hierarchical linear mixed models. *Statistica Sinica*, 11(1):77–95, 2001.
- V. H. Swallow und J. F. Monahan. Monte Carlo comparison of ANOVA, MIVQUE, REML and ML estimators of variance components. *Technometrics*, 26(1):47–57, 1984.
- T. J. Sweeting. Uniform asymptotic normality of the maximum likelihood estimator. *Annals of Statistics*, 8(6):1375–1381, 1980.
- L. Tierney. Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, 22(4):1701–1728, 1994.
- E. Wagner. *Credit Default Swaps und Informationsgehalt*. Inauguraldissertation, Universität Linz, 2007.
- R. A. Wijsman. Invariant measures on groups and their use in statistics. *Lecture Notes-Monograph Series*, 14(1), 1990. Hayward, CA: Institute of Mathematical Statistics.
- A. Zeileis und T. Hothorn. Diagnostic Checking in Regression Relationships. *R News*, 2:7–10, 2002.
- B. Zhang, H. Zhou und H. Zhu. Explaining credit default swap spreads with the equity volatility and jump risks of individual firms. *Review of Financial Studies*, 22(12):5099–5131, 2009.
- D. Zhang und M. Davidian. Linear Mixed Models with Flexible Distributions of Random Effects for Longitudinal Data. *Biometrics*, 57(3):795–802, 2001.
- X. Zhang, B. Schwaab und A. Lucas. Conditional probabilities and contagion measures for Euro Area sovereign default risk. Working paper, 2012. URL [http://www.berndschwaab.eu/papers/EAsovereigns\\_Feb12\\_.pdf](http://www.berndschwaab.eu/papers/EAsovereigns_Feb12_.pdf).

*Literatur*

- Y. Zhao, J. Staudenmayer, B. A. Coull und M. P. Wand. General Design Bayesian Generalized Linear Mixed Models. *Statistical Science*, 21(1):35–51, 2006.
- D. L. Zimmerman und J. F. Monahan. Computationally efficient restricted maximum likelihood estimation of generalized covariance funtions. *Mathematical Geology*, 21(7):655–672, 1989.