

Ein Score-Test auf Messfehler in Ratingmigrationen

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften

Der Fakultät Statistik
der Technischen Universität Dortmund
vorgelegt von

Sebastian Voß

Dortmund, den 14. März 2012

Inhaltsverzeichnis

1	Einleitung	1
2	Modell ohne Messfehler	4
2.1	Multiplikatives Intensitätsmodell	4
2.2	Parameterschätzung	6
2.3	Konsistenz der Parameterschätzer	9
3	Messfehler in den Ratingdaten	18
3.1	Modellierung von Messfehlern	18
3.2	Folgen der Vernachlässigung	20
4	Score-Test auf Messfehler	23
4.1	Herleitung des Tests	23
4.2	Asymptotische Verteilung unter Hypothese	29
4.3	Güte und Niveau in endlichen Stichproben	41
5	Empirische Analyse von Ratingwechselintensitäten und -wahrscheinlichkeiten	46
5.1	Daten	46
5.2	Modellanpassung	49
5.3	Bestimmung von Ratingwechselwahrscheinlichkeiten	55
5.4	Testergebnisse im Ratingmodell	57
6	Zusammenfassung und Fazit	58
	R-Code	61
	Literaturverzeichnis	68

1 Einleitung

Die Eigenkapitalvorschriften des Baseler Ausschuss für Bankenaufsicht (2004) „Basel II“ legen Banken der Europäischen Union nahe, das zur Absicherung des Kreditrisikos benötigte Eigenkapital mit Hilfe von Kreditportfoliomodellen zu bestimmen. Wichtige Inputfaktoren dieser Modelle sind Ratingwechselwahrscheinlichkeiten und insbesondere Ausfallwahrscheinlichkeiten von Schuldern, welche sich in Migrationsmatrizen zusammenfassen lassen. Basel II schreibt den Banken vor, diese Migrationsmatrizen mit Hilfe von Datenhistorien, welche mindestens die letzten fünf Jahre umfassen, zu schätzen (vgl. Baseler Ausschuss für Bankenaufsicht (2004), S. 66). Nach Kadam und Lenk (2008) führt die Annahme der Homogenität aller Schuldner dabei zu einer Fehleinschätzung des Kreditrisikos. Schuldnerspezifische Eigenschaften müssen demnach bei der Bestimmung der Wechselwahrscheinlichkeiten in Form von Kovariablen berücksichtigt werden.

Mögliche Ansätze zur Bestimmung von Migrationsmatrizen in Abhängigkeit von Kovariablen liefern zum Beispiel Kim (1999), Nickell et al. (2000), Koopman et al. (2009) und Koopman et al. (2011). Die genannten Autoren berücksichtigen dabei unter anderem Wirtschaftsindikatoren in ihren Modellen und gelangen zu dem Resultat, dass die Makroökonomie ebenfalls einen Einfluss auf die Wechselwahrscheinlichkeiten der Schuldner hat. Kim (1999) und Nickell et al. (2000) wählen einen zeitdiskreten Ansatz über vierteljährliche bzw. einjährige Probit-Modelle und berücksichtigen somit nicht, dass sich die Bonität - und damit auch das Rating - eines Schuldners zu jeder Zeit ändern kann. Die Möglichkeit von mehrfachen Ratingwechseln innerhalb des Zeithorizonts, sowie die Dauer des Aufenthaltes in einer Ratingstufe wird bei der Modellierung vernachlässigt. Koopman et al. (2009) und Koopman et al. (2011) berücksichtigen den zeitstetigen Charakter der Schuldnerbonität bei der Bestimmung der Wechselwahrscheinlichkeiten und modellieren das Rating als Markov-Prozess. Die Berechnung der Migrationsmatrizen basiert in diesen Arbeiten auf den Intensitäten der Ratingwechsel. Die Schätzung dieser Intensitäten gelingt Koopman et al. (2009) und Koopman et al. (2011) mit Hilfe von bayesianischen Methoden. Alle bislang genannten Arbeiten verwenden Daten der Ratingagenturen Moody's und Standard & Poor's zur Schätzung ihrer Modelle. Banken verwenden allerdings üblicherweise interne Ratingsysteme zur Bonitätsbewertung ihrer Schuldner (IRB-Ansatz), da insbesondere kleinere Schuldnerunternehmen nicht durch die externen Ratingagenturen bewertet wer-

den. Diese Schuldner müssten mit dem vom Baseler Ausschuss für Bankenaufsicht (2004) vorgeschriebenen Standardsatz durch Eigenkapital abgesichert werden. Da dieser jedoch recht konservativ bemessen ist, führt der IRB-Ansatz zu einer deutlichen Kostenersparnis. Der zeitstetige Charakter des Ratings erfordert eine zumindest annähernd kontinuierliche Überwachung der Schuldnerbonitäten. Nach Mählmann (2006) ist diese Voraussetzung jedoch bei bankinternen Ratingsystemen nicht zwangsläufig erfüllt. Werden die Ratings nur in unregelmäßigen Abständen überprüft, kann es zu Messfehlern bei der Erfassung der Ratingwechselzeitpunkte kommen.

Messfehler in den Migrationszeitpunkten haben eine Verzerrung der geschätzten Intensitäten - und damit auch der Migrationsmatrizen - zur Folge und sollten demnach unbedingt im Modell berücksichtigt werden. Dies ist jedoch mit einem nicht unerheblichen Aufwand verbunden und bedarf einer recht genauen Vorstellung über die Struktur des Messfehlers. Es bietet sich daher an, zunächst mit Hilfe eines statistischen Tests zu untersuchen, ob Messfehler in den Ratingdaten vorhanden sind. Ein solcher Test wird in der folgenden Arbeit - basierend auf einer Idee von Lancaster (1990) - entwickelt. Als Grundlage dient dabei ein multiplikatives Intensitätsmodell, ähnlich dem von Koopman et al. (2009) und Koopman et al. (2011). Die zugrunde liegende Zeitachse ist in diesem Fall jedoch nicht die Kalenderzeit, sondern das Alter der Schuldner, d.h. die vergangene Zeit seit Aufnahme des Kredits. Dies hat nach Weißbach und Walter (2010) den Vorteil, dass sich die Schuldner als unabhängig auffassen lassen. Koopman et al. (2009) und Koopman et al. (2011) sind aufgrund der Kalenderzeitachse gezwungen, die Abhängigkeit der Schuldner durch einen latenten Frailty-Prozess innerhalb der Intensitäten zu berücksichtigen, was die Schätzung der Modellparameter deutlich erschwert.

Das dem Test zu Grunde liegende Intensitätsmodell enthält ebenfalls einen Frailty-Prozess, welcher in diesem Fall jedoch der Modellierung der Messfehler dient. Der unbeobachtbare Messfehlerprozess hat den Erwartungswert 1 und beeinflusst die Intensität multiplikativ. Von Interesse ist, ob sich die Volatilität dieses Prozesses signifikant von Null unterscheidet. Ist dies nicht der Fall, nimmt der Messfehlerprozess konstant den Erwartungswert 1 an und existiert somit praktisch nicht. Die Likelihood dieses Intensitätsmodells ist aufgrund der Messfehler allerdings nicht beobachtbar. Dieses Problem lässt sich lösen, indem der Erwartungswert der Likelihood bezüglich der Messfehler gebildet wird. Um dabei einer Verteilungsannahme aus dem Weg zu gehen, wird die Likelihood vor der Erwartungswertbildung durch eine Taylorapproximation ersetzt. Diese Approximation entspricht bei geringen Messfehlern einer guten Annäherung an die tatsächliche Likelihood und hat den Vorteil, dass für die Bestimmung des Erwartungswerts lediglich die ersten beiden Momente der Messfehlerverteilung benötigt werden. Ergebnis dieser Prozedur ist eine Likelihood, welche neben den zu schätzenden Modellparametern einen Messfehler-

varianzparameter enthält und sich somit als Grundlage für das vorliegende Testproblem eignet. Da überprüft werden soll, ob sich die Varianz der Messfehler signifikant von Null unterscheidet, liegt der zu testende Parameter unter der Nullhypothese am Rand des zulässigen Parameterbereichs, was zu unschönen Eigenschaften des entsprechenden ML-Schätzers führt. Um der Schätzung des Varianzparameters aus dem Weg zu gehen, wird in der vorliegenden Arbeit ein Score-Test auf Basis der erwarteten approximierten Likelihood konstruiert. Dieses Testverfahren hat den Vorteil, dass lediglich die Modellparameter zu schätzen sind, die nicht durch die Nullhypothese festgelegt werden.

Datengrundlage für den empirischen Teil der vorliegenden Arbeit sind Ratinghistorien aus dem Geschäftskundenportfolio einer großen deutschen Bank aus dem Zeitraum von 1997 bis 2007. Das multiplikative Intensitätsmodell wird unter Berücksichtigung von Branche und Standort des Schuldners, sowie den Wirtschaftsindikatoren Bruttoinlandsprodukt, Industrieproduktion, Interbankzinssätzen und Aktienindexvolatilität angepasst. Die Anwendung des entwickelten Score-Tests auf das geschätzte Modell bestätigt den Verdacht bezüglich der Existenz von Messfehlern in bankinternen Ratingsystemen.

Die Arbeit ist wie folgt aufgebaut: Zunächst wird in Kapitel 2 das Intensitätsmodell ohne Berücksichtigung von Messfehlern und die damit einhergehende Parameterschätzung im Detail vorgestellt. Anschließend geht Kapitel 3 auf die Messfehlerproblematik in den Ratingdaten ein. Dort werden Methoden zur Berücksichtigung erläutert, sowie die Konsequenzen der Vernachlässigung anhand einer Simulationsstudie aufgezeigt. Kapitel 4 stellt den Kern der Arbeit dar. Hier wird ein Score-Test auf Messfehler in den Ratingwechselzeitpunkten entwickelt. Die Herleitung der Teststatistik erfolgt dabei in Kapitel 4.1, während Kapitel 4.2 den Nachweis der asymptotischen Verteilung unter der Nullhypothese liefert. Abgeschlossen wird Kapitel 4 mit einigen Simulationsstudien zur Güte des entwickelten Tests in endlichen Stichproben. In Kapitel 5 werden zunächst die vorliegenden Daten im Detail vorgestellt. Es folgt eine Anpassung des Intensitätsmodell aus Kapitel 2 und darauf aufbauend die Bestimmung von Ein-Jahres-Migrationsmatrizen für verschiedene Ausprägungen der verwendeten Kovariablen. Das Kapitel endet mit der Anwendung des entwickelten Tests auf das geschätzte Modell.

2 Modell ohne Messfehler

Das Rating eines Schuldners lässt sich als zeitstetiger Markov Prozess mit endlichem Zustandsraum auffassen. Die Modellierung dieses Prozesses erfolgt über die Intensitäten der möglichen Ratingwechsel, welche als Grundlage für die Schätzung von Migrationsmatrizen dienen. Das Ratingsystem, das den vorliegenden Daten (Näheres dazu in Kapitel 5.1) zu Grunde liegt, umfasst 21 Ratingklassen. Dies entspricht in etwa dem Standard der drei größten Ratingagenturen Moody's (21), Standard & Poor's (20) und Fitch (22). Unter der Annahme, dass die Ratingklasse 21 - welche für die Insolvenz eines Schuldners steht - absorbierend ist, sind theoretisch insgesamt 400 verschiedene Ratingwechsel möglich, deren Intensitäten es zu schätzen gilt. Die in der Praxis üblichen Stichprobengrößen sind jedoch bei weitem nicht groß genug, um die dazu notwendige Anzahl an Modellparametern zu bestimmen, da gerade Ratingwechsel in weit entfernte Klassen nur äußerst selten zu beobachten sind. Der überwiegende Teil der Wechsel erfolgt in die benachbarte Ratingklasse. Unter der Annahme, dass es sich bei der Bonität eines Schuldners um einen im Zeitablauf stetigen Prozess handelt und die Bank auf Bonitätsveränderung unmittelbar mit einer Ratinganpassung reagieren kann, sollten Ratingwechsel um mehr als eine Stufe theoretisch sogar unmöglich sein. In der Praxis ist dies aufgrund von Verzögerungen im Informationsfluss zwischen Bank und Schuldner allerdings nicht immer gewährleistet, so dass in den Daten mitunter auch Ratingwechsel höherer Ordnung auftauchen (vgl. Tabelle 5.1). In Anlehnung an Weißbach und Mollenhauer (2011) sollen diese Ratingwechsel jedoch im Folgenden vernachlässigt werden, d.h. die Modellierung und Schätzung der Intensitäten beschränkt sich auf die Wechsel in die benachbarte Ratingklasse. Die übrigen Intensitäten werden als konstant 0 angenommen. Dies impliziert, dass Ratingwechsel um mehr als eine Stufe nur durch eine Folge einstufiger Sprünge zustande kommen können und nicht direkt möglich sind.

2.1 Multiplikatives Intensitätsmodell

Sei τ die Länge des Zeitraums, über den die Schuldner beobachtet werden und sei $(R_i(t), t \in [0, \tau])$ ein zeitstetiger Markov Prozess, welcher für $i = 1, \dots, n$ die Ratingklasse des Schuldners i zum Zeitpunkt t beschreibt. Der Zustandsraum entspricht somit den Ratingklassen

$\{1, \dots, D\}$. Die Intensität des Ratingwechsel von p nach q sei für den Fall $|p - q| = 1$ und $p < D$ gegeben durch

$$\lambda_{0(p-q)} \exp\left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(t)\right). \quad (2.1)$$

Die Intensitäten der übrigen Ratingwechsel werden als konstant 0 angenommen. $(\mathbf{X}_i(t), t \in [0, \tau])$ bezeichnet den Kovariablenprozess von Schuldner i , welcher sowohl zeitabhängige, als auch zeitkonstante Komponenten enthalten kann. Die Prozesse seien linksstetig mit existierenden rechtsseitigen Grenzwerten, sowie fast sicher beschränkt und für $i = 1, \dots, n$ unabhängig identisch verteilt. Die Kovariableneinflüsse werden durch $\boldsymbol{\beta}_{(p-q)} \in \mathbb{R}^d$ beschrieben, während $\lambda_{0(p-q)} \in \mathbb{R}^+$ eine konstante Basishazardfunktion darstellt. Im Modell (2.1) wird unterstellt, dass in einem hinreichend kleinen Zeitraum lediglich Ratingwechsel um genau eine Stufe in den Ausgangsratings 1 bis 20 mit positiver Wahrscheinlichkeit auftreten können und gleichgerichtete Ratingwechsel - unabhängig vom Ausgangsrating - die gleiche Intensität aufweisen. Außerdem wird die mit D codierte Ratingklasse als absorbierend aufgefasst. Die Zeit t entspricht nicht der Kalenderzeit, sondern dem Alter eines Schuldners (in Tagen). Der Zeitpunkt $t = 0$ stellt also den Start der Kreditlaufzeit dar. Auf diese Weise ist sichergestellt, dass den Ratings $R_1(t), \dots, R_n(t)$ zu jedem Zeitpunkt t unterschiedliche makroökonomische oder bankinterne Bedingungen zu Grunde liegen.¹ Die Ratings und die Kovariablenprozesse lassen sich demnach als unabhängig auffassen.

Da nicht davon auszugehen ist, dass das Rating eines Schuldners zum Zeitpunkt $t = 0$ in irgendeiner Form von den Parametern der Intensitäten abhängt, lässt sich die für die Parameterschätzung relevante Information der Ratinghistorie von Schuldner i alternativ durch die rechtsstetigen Zählprozesse $(N_{pqi}(t), t \in [0, \tau])$ und die linksstetigen Risikoprozesse $(Y_{pi}(t), t \in [0, \tau])$ beschreiben, wobei $p = 1, \dots, D$ mit $p \neq q$. Der Prozess $Y_{pi}(t) \in \{0, 1\}$ nimmt den Wert 1 an, wenn sich Schuldner i unmittelbar vor dem Zeitpunkt t in der Ratingklasse p befindet und somit unter dem Risiko steht diese zu verlassen. $N_{pqi}(t)$ zählt die Anzahl der bis zum Zeitpunkt t an Schuldner i beobachteten Ratingwechsel von Klasse p nach Klasse q . Die Risikoprozesse werden als unabhängig und identisch verteilt angenommen. Für die Intensität von $(N_{pqi}(t), t \in [0, \tau])$ gilt wegen (2.1)

$$\lambda_{pqi}(t) = \begin{cases} Y_{pi}(t) \lambda_{0(p-q)} \exp\left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(t)\right) & \text{falls } p < D \text{ und } |p - q| = 1 \\ 0 & \text{sonst} \end{cases}. \quad (2.2)$$

¹Die Makroökonomie kann zwar teilweise in Form von Kovariablen berücksichtigt werden, es ist jedoch nicht davon auszugehen, dass sich auf diese Weise sämtliche relevanten wirtschaftliche Einflüsse erfassen lassen.

2.2 Parameterschätzung

Die Modellparameter lassen sich anhand der beobachteten einstufigen Ratingwechsel und der Pfade der Risikoprozesse mit Hilfe der Maximum-Likelihood Methode schätzen. Sei $\boldsymbol{\theta} = (\lambda_{0(+1)}, \boldsymbol{\beta}_{+1}^\top, \lambda_{0(-1)}, \boldsymbol{\beta}_{-1}^\top)^\top$ der Vektor der Modellparameter. Nach Andersen et al. (1993), S. 585, ist die Likelihood für einen Pfad der Zählprozesse von Schuldner i mit der Intensität (2.2) in diesem Fall gegeben durch

$$L_i(\boldsymbol{\theta}) = \prod_{p,q:p \neq q} \left(\prod_{t \in [0,\tau]} \lambda_{pqi}(t)^{dN_{pqi}(t)} \right) \exp \left(- \int_0^\tau \lambda_{pqi}(t) dt \right). \quad (2.3)$$

Dabei wird angenommen, dass der Zensierungsmechanismus unabhängig von den Zählprozessen $N_{pqi}(t)$ und somit uninformativ ist. $dN_{pqi}(t)$ bezeichnet die Änderung des Zählprozesses N_{pqi} zum Zeitpunkt t , d.h. $dN_{pqi}(t)$ nimmt lediglich die Werte 0 und 1 an. Sei $N_i(t)$ die Anzahl der bis zum Zeitpunkt t an Schuldner i beobachteten Ratingwechsel und seien $T_{i1}, \dots, T_{iN_i(\tau)}$ die Migrationszeitpunkte im Beobachtungszeitraum. Dann gilt:

$$L_i(\boldsymbol{\theta}) = \prod_{p,q:p \neq q} \left(\prod_{j=1}^{N_i(\tau)} \lambda_{pqi}(T_{ij})^{dN_{pqi}(T_{ij})} \right) \exp \left(- \int_0^\tau \lambda_{pqi}(t) dt \right). \quad (2.4)$$

Ratingwechsel um mehr als eine Stufe haben im Modell (2.2) die Intensität 0. Die Intensitäten der einstufigen Ratingupgrades bzw. -downgrades unterscheiden sich lediglich durch den Risikoprozess Y_{pi} , welcher vom Rating eines Schuldners i zum Zeitpunkt t abhängt. Somit lässt sich (2.4) umformen zu

$$\begin{aligned} L_i(\boldsymbol{\theta}) &= \prod_{p,q:|p-q|=1} \prod_{j=1}^{N_i(\tau)} \left(Y_{pi}(T_{ij}) \lambda_{0(p-q)} \exp \left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(T_{ij}) \right) \right)^{dN_{pqi}(T_{ij})} \\ &\quad \times \prod_{p,q:|p-q|=1} \exp \left(- \int_0^\tau \lambda_{pqi}(t) dt \right) \\ &= \prod_{p,q:|p-q|=1} \prod_{j=1}^{N_i(\tau)} \left(\lambda_{0(p-q)} \exp \left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(T_{ij}) \right) \right)^{dN_{pqi}(T_{ij})} \\ &\quad \times \exp \left(- \sum_{p,q:p-q=-1} \int_0^\tau Y_{pi}(t) \lambda_{0(p-q)} \exp \left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(t) \right) dt \right) \\ &\quad \times \exp \left(- \sum_{p,q:p-q=+1} \int_0^\tau Y_{pi}(t) \lambda_{0(p-q)} \exp \left(\boldsymbol{\beta}_{(p-q)}^\top \mathbf{X}_i(t) \right) dt \right). \end{aligned}$$

Damit folgt

$$\begin{aligned}
 L_i(\boldsymbol{\theta}) &= \prod_{j=1}^{N_i(\tau)} \left(\lambda_{0(-1)} \exp \left(\boldsymbol{\beta}_{-1}^\top \mathbf{X}_i(T_{ij}) \right) \right)^{d \sum_{p,q:p-q=-1} N_{pqi}(T_{ij})} \\
 &\quad \times \exp \left(- \int_0^\tau \left(\sum_{p,q:p-q=-1} Y_{pi}(t) \right) \lambda_{0(-1)} \exp \left(\boldsymbol{\beta}_{-1}^\top \mathbf{X}_i(t) \right) dt \right) \\
 &\quad \times \prod_{j=1}^{N_i(\tau)} \left(\lambda_{0(+1)} \exp \left(\boldsymbol{\beta}_{+1}^\top \mathbf{X}_i(T_{ij}) \right) \right)^{d \sum_{p,q:p-q=+1} N_{pqi}(T_{ij})} \\
 &\quad \times \exp \left(- \int_0^\tau \left(\sum_{p,q:p-q=+1} Y_{pi}(t) \right) \lambda_{0(+1)} \exp \left(\boldsymbol{\beta}_{+1}^\top \mathbf{X}_i(t) \right) dt \right). \quad (2.5)
 \end{aligned}$$

Sei nun

$$N_{hi}(t) := \sum_{p,q:p-q=h} N_{pqi}(t), \quad h \in \{-1, +1\} \quad (2.6)$$

die Anzahl der an Schuldner i beobachteten einstufigen Upgrades ($h = +1$) bzw. Downgrades ($h = -1$) zum Zeitpunkt t . Der Prozess

$$Y_{hi}(t) := \sum_{p,q:p-q=h} Y_{pi}(t), \quad h \in \{-1, +1\}, \quad (2.7)$$

gibt an, ob ein Schuldner zum Zeitpunkt t unter dem Risiko eines einstufigen Ratingwechsels steht. Man beachte, dass Y_{-1i} und Y_{+1i} im Allgemeinen nicht gleich sind. Während $Y_{-1i}(t)$ bis zum Ende der Beobachtungsdauer des Schuldners den Wert 1 aufweist und erst dann den Wert 0 annimmt, kann der Prozess Y_{+1i} auch schon vor dem Ende der Beobachtungsdauer zwischenzeitlich den Wert 0 annehmen. Dies ist genau dann der Fall, wenn Schuldner i in die höchste Ratingklasse aufsteigt. Solange sich sein Rating nicht wieder verschlechtert, besteht für ihn nicht mehr das Risiko eines einstufigen Upgrades, da er sich nicht weiter verbessern kann. In Formel (2.7) äußert sich das dadurch, dass für $p = 1$ kein q mit $p - q = +1$ existiert und $Y_{+1i}(t)$ somit einen Summanden weniger aufweist, d.h. es gilt

$$Y_{-1i}(t) = \sum_{p=1}^{D-1} Y_{pi}(t) \quad \text{und} \quad Y_{+1i}(t) = \sum_{p=2}^{D-1} Y_{pi}(t). \quad (2.8)$$

Seien $\{T_{hi1}, \dots, T_{hiN_{hi}(\tau)}\} \subset \{T_{i1}, \dots, T_{iN_i(\tau)}\}$, $h \in \{-1, +1\}$, die $N_{hi}(\tau)$ an Schuldner i beobachteten einstufigen Upgrade- ($h = +1$) bzw. Downgradezeitpunkte ($h = -1$). Da $dN_{hi}(t)$ an den Stellen $T_{hi1}, \dots, T_{hiN_{hi}(\tau)}$ stets den Wert 1 annimmt, ist eine alternative

Darstellung von (2.5) gegeben durch

$$\begin{aligned}
 L_i(\boldsymbol{\theta}) &= \prod_{h \in \{-1, +1\}} \prod_{j=1}^{N_{hi}(\tau)} \left(\lambda_{0h} \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij}) \right) \right) \\
 &\quad \times \exp \left(- \int_0^\tau Y_{hi}(t) \lambda_{0h} \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t) \right) dt \right) \\
 &=: \prod_{h \in \{-1, +1\}} L_{hi}(\boldsymbol{\theta}_h),
 \end{aligned} \tag{2.9}$$

wobei $\boldsymbol{\theta}_h = (\lambda_{0h}, \boldsymbol{\beta}_h^\top)^\top$, $h \in \{-1, +1\}$. Aufgrund der Tatsache, dass die Intensitäten gleichgerichteter einstufiger Ratingwechsel dieselben Parameter aufweisen, sind die Ausgangs- und Zielratings der einzelnen Migrationen eines Schuldners für die Inferenz also nicht relevant. Anstelle der Zählprozesse $N_{pqi}(t)$, $p = 1, \dots, D-1$, $q = 1, \dots, D$ mit Intensitäten (2.2) genügt es also für jeden Schuldner die beiden Zählprozesse $N_{-1i}(t)$ und $N_{+1i}(t)$ zu betrachten. Diese haben nach (2.9) die Intensität

$$\lambda_{hi}(t) = Y_{hi}(t) \lambda_{0h} \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t) \right), \quad h \in \{-1, +1\}. \tag{2.10}$$

Aufgrund der Argumentation aus Abschnitt 2.1 lassen sich die Ratinghistorien der einzelnen Schuldner als unabhängig auffassen. Damit ergibt sich die Likelihood für n Schuldner zu

$$L(\boldsymbol{\theta}) = \prod_{h \in \{-1, +1\}} \prod_{i=1}^n L_{hi}(\boldsymbol{\theta}_h), \tag{2.11}$$

so dass die log-Likelihood für die Parameterinferenz gegeben ist durch

$$\begin{aligned}
 \log L(\boldsymbol{\theta}) &= \sum_{h \in \{-1, +1\}} \sum_{i=1}^n \left(\sum_{j=1}^{N_{hi}(\tau)} \left(\log(\lambda_{0h}) + \boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij}) \right) \right. \\
 &\quad \left. - \int_0^\tau Y_{hi}(t) \lambda_{0h} \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t) \right) dt \right).
 \end{aligned} \tag{2.12}$$

Aufgrund der in (2.11) dargestellten Faktorisierung der Likelihood können die Upgrade- und Downgradeparameter unabhängig voneinander geschätzt werden indem die beiden Faktoren getrennt voneinander bezüglich $\boldsymbol{\theta}_{+1}$ und $\boldsymbol{\theta}_{-1}$ maximiert werden. Eine geschlossene Form des Maximums existiert nicht, so dass dazu auf numerische Verfahren zurückgegriffen werden muss. Liegen zeitabhängige Kovariablen vor, lassen sich die Integrale in (2.12) nicht analytisch lösen, so dass die numerische Bestimmung des Maximums deutlich aufwendiger wird. In diesem Fall bietet sich eine Parameterschätzung in zwei Stufen an, wobei zunächst nur die Kovariableneinflüsse $\hat{\boldsymbol{\beta}}_h$ mit der partiellen Likelihood von Andersen und Gill (1982) bestimmt werden. Diese ergibt sich aus einer Faktorisierung der

vollständigen Likelihood und ist gegeben durch

$$\begin{aligned} L^P(\boldsymbol{\beta}_{(+1)}, \boldsymbol{\beta}_{(-1)}) &= \prod_{h \in \{+1, -1\}} L_h^P(\boldsymbol{\beta}_h) \\ &= \prod_{h \in \{+1, -1\}} \prod_{i=1}^n \prod_{j=1}^{N_{hi}(\tau)} \frac{\exp(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij}))}{\sum_{k=1}^n Y_{hk}(T_{hij}) \exp(\boldsymbol{\beta}_h^\top \mathbf{X}_k(T_{hij}))} \end{aligned} \quad (2.13)$$

Dabei handelt es sich um eine Verallgemeinerung der partiellen Likelihood von Cox (1972) aus der Überlebenszeitanalyse. Unter der Annahme, dass es sich bei den Risikoprozessen Y_{-1i} und Y_{+1i} , $i = 1, \dots, n$, um linksstetige Prozesse handelt, entspricht jeder Faktor von (2.13) der bedingten Wahrscheinlichkeit für den beobachteten Ratingwechsel, gegeben, dass eines der unter Risiko stehenden Unternehmen das Rating wechseln muss. Die Maximierung von $\log L^P$ erfolgt ebenfalls numerisch. Hier lässt sich allerdings auf in den gängigen Softwarepaketen implementierte Methoden zurückgreifen, welche erfahrungsgemäß schnelle und zuverlässige Ergebnisse liefern.

Einen Schätzer für λ_{0h} erhält man schließlich, indem man den mit Hilfe von $\log L^P$ ermittelten Schätzer $\hat{\boldsymbol{\beta}}_h$ in (2.12) einsetzt. Mit

$$\frac{\partial}{\partial \lambda_{0h}} \log L_h(\lambda_{0h}, \hat{\boldsymbol{\beta}}_h) = \sum_{i=1}^n \left(\sum_{j=1}^{N_{hi}(\tau)} \frac{1}{\lambda_{0h}} - \int_0^\tau Y_{hi}(t) \exp(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)) dt \right)$$

ergibt sich

$$\hat{\lambda}_{0h} = \frac{\sum_{i=1}^n N_{hi}(\tau)}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)) dt}. \quad (2.14)$$

2.3 Konsistenz der Parameterschätzer

In diesem Unterkapitel soll nun gezeigt werden, dass die vorgestellten Schätzer unter schwachen Regularitätsannahmen konsistent für die wahren Modellparameter sind. Der Nachweis erfolgt mit Hilfe von Martingalgrenzwerttheorie.

Die multivariaten Prozesse $(N_{+1i}(t), N_{-1i}(t), Y_{+1i}(t), Y_{-1i}(t), \mathbf{X}_i(t), t \in [0, \tau])$ auf dem Wahrscheinlichkeitsraum $(\Omega^{(n)}, \mathcal{F}^{(n)}, \mathbb{P}^{(n)})$ seien an die natürliche Filtration $(\mathcal{F}_t^{(n)}, t \in [0, \tau])$ adaptiert. Dadurch ist u.a. gewährleistet, dass die Intensität von $N_{hi}(t)$ vorhersagbar ist (vgl. Andersen et al. (1993), S. 75). Im weiteren Verlauf der Arbeit bezeichnen $\lambda_{0h}^{(0)}$ und $\boldsymbol{\beta}_h^{(0)}$ die wahren Parameter, während λ_{0h} und $\boldsymbol{\beta}_h$ beliebige Werte aus den jeweiligen Parameterräumen darstellen. Es gelten die folgenden Regularitätsannahmen:

(A1) $P(N_{hi}(\tau) < \infty) = 1$

(A2) $P(Y_{hi}(t) = 1 \forall t \in [0, \tau]) > 0$ für alle $h \in \{+1, -1\}$, $i = 1, \dots, n$.

(A3) Die Fisher-Informationsmatrix \mathcal{I}^P von L^P aus Gleichung (2.13) ist positiv definit.

(A4) $\lambda_{0h}^{(0)} \in (c_h^l, c_h^u)$, wobei $0 < c_h^l < c_h^u < \infty$

(A5) Es existiert eine kompakte Umgebung \mathcal{B}_h um $\beta_h^{(0)}$ und auf $[0, \tau]$ beschränkte, deterministische Funktionen $\mu_h(t, \beta_h)$, so dass

$$\sup_{t \in [0, \tau]} \left| \frac{1}{n} \sum_{i=1}^n Y_{hi}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) - \mu_h(t, \beta_h) \right| \xrightarrow{\mathcal{P}} 0$$

für alle $\beta_h \in \mathcal{B}_h$. $\mu_h(t, \beta_h)$ sei dabei für alle $t \in [0, \tau]$ stetig in β_h und

$$\int_0^\tau \mu(t, \beta_h) dt \neq 0 \quad \forall \beta_h \in \mathcal{B}_h.$$

Die Annahmen sind nicht sehr streng. (A1) besagt, dass in endlicher Zeit auch nur eine endliche Anzahl an Ratingwechseln möglich ist, während Annahme (A2) voraussetzt, dass jeder Schuldner mit positiver Wahrscheinlichkeit über den gesamten Beobachtungszeitraum unter dem Risiko eines Ratingwechsels steht. Annahme (A3) gewährleistet die Invertierbarkeit der Fisher-Informationsmatrix der partiellen Likelihood. Da \mathcal{I}^P sich aus den Erwartungswerten der zweiten Ableitungen von $\log L^P$ ergibt, ist die Matrix symmetrisch und damit zumindest positiv semidefinit. Positiv definit ist sie genau dann, wenn sie regulär ist. Da alle zweiten Ableitungen von $\log L^P$ dieselbe Struktur aufweisen, besteht eine notwendige Bedingung für Nichtregularität darin, dass ein perfekter funktionaler Zusammenhang zwischen den Kovariablenmomenten besteht. Dies lässt sich jedoch in der Praxis ruhigen Gewissens ausschließen. Annahme (A4) stellt sicher, dass die Basisintensitäten der Ratingwechsel nicht beliebig groß oder beliebig klein sein können, während Annahme (A5) im Wesentlichen besagt, dass der zufällige, in β_h stetige Intensitätsprozess über den gesamten Beobachtungszeitraum gleichmäßig gegen eine ebenfalls in β_h stetige erwartete Intensität konvergiert.

Da die Modellparameter nacheinander in zwei Schritten geschätzt werden, gilt es zunächst, die Konsistenz der Schätzer für die Kovariableneinflüsse $\beta_h^{(0)}$ nachzuweisen.

Satz 2.1 *Die Annahmen (A1) bis (A4) seien erfüllt. Dann ist $\hat{\beta} = (\hat{\beta}_{+1}^\top, \hat{\beta}_{-1}^\top)^\top$ aus Unterkapitel 2.2 ein konsistenter und asymptotisch normalverteilter Schätzer für $\beta^{(0)} = (\beta_{+1}^{(0)\top}, \beta_{-1}^{(0)\top})^\top$, d.h. es gilt*

$$\sqrt{n} (\hat{\beta} - \beta^{(0)}) \xrightarrow{\mathcal{D}} N(\mathbf{0}, (\mathcal{I}^P)^{-1}).$$

Beweis: Andersen und Gill (1982) weisen die Behauptung mit Hilfe von Martingalgrenzwerttheorie nach (Satz 3.2). Um das Resultat der beiden Autoren anwenden zu können, müssen lediglich zwei Bedingungen überprüft werden:

1.

$$\int_0^\tau \lambda_{0h}^{(0)} dt < \infty \quad (2.15)$$

2. Es existiert eine Umgebung \mathcal{B}_h um $\beta_h^{(0)}$, so dass

$$\mathbb{E} \left[\sup_{t \in [0, \tau], \beta_h \in \mathcal{B}_h} Y_{hi}(t) \|\mathbf{X}_i(t)\|^2 \exp \left(\beta_h^\top \mathbf{X}_i(t) \right) \right] < \infty \quad (2.16)$$

Die Bedingung (2.15) folgt dabei unmittelbar aus **(A4)**. Die Gültigkeit von (2.16) lässt sich durch die Beschränktheit der Kovariablenprozesse rechtfertigen. Dies hat zur Folge, dass der Ausdruck innerhalb des Erwartungswerts aus (2.16) fast sicher beschränkt ist und damit auch der Erwartungswert selbst. Zusammen mit **(A1)** bis **(A3)** sind nach Andersen und Gill (1982), Satz 4.1, sämtliche Bedingungen von Andersen und Gill (1982), Satz 3.2 erfüllt. \square

Der Schätzer für $\lambda_{0h}^{(0)}$ ist nach Gleichung (2.14) gegeben durch

$$\hat{\lambda}_{0h}(\hat{\beta}_h) = \frac{\sum_{i=1}^n N_{hi}(\tau)}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp \left(\hat{\beta}_h^\top \mathbf{X}_i(t) \right) dt}$$

und hängt somit von dem Schätzer für die Kovariableneinflüsse ab. Zunächst lässt sich nachweisen, dass $\hat{\lambda}_{0h}(\beta_h)$ punktweise gegen einen Wert $\lambda_{0h}(\beta_h)$ konvergiert (Satz 2.2). Anschließend liefert Satz 2.3 die Erkenntnis, dass $\lambda_{0h}(\beta_h^{(0)}) = \lambda_{0h}^{(0)}$ ist. Auf Grundlage dieser beiden Resultate lässt sich schließlich in Satz 2.4 die Konsistenz von $\hat{\lambda}_{0h}(\hat{\beta}_h)$ zeigen.

Satz 2.2 Die Annahmen **(A1)**, **(A4)** und **(A5)** seien erfüllt. Sei \mathcal{B}_h die kompakte Umgebung um $\beta_h^{(0)}$ aus **(A5)**. Dann gilt

$$\hat{\lambda}_{0h}(\beta_h) \xrightarrow{\mathcal{P}} \lambda_{0h}(\beta_h) \quad \forall \beta_h \in \mathcal{B}_h.$$

$\lambda_{0h}(\beta_h)$ ist dabei stetig in β_h .

Beweis: Wegen **(A1)** gilt nach dem Gesetz der großen Zahl

$$\frac{1}{n} \sum_{i=1}^n N_{hi}(\tau) \xrightarrow{\mathcal{P}} \mathbb{E} [N_{h1}(\tau)] < \infty. \quad (2.17)$$

Aufgrund von **(A5)** gilt außerdem für alle $\beta_h \in \mathcal{B}_h$

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp\left(\beta_h^\top \mathbf{X}_i(t)\right) dt \xrightarrow{\mathcal{P}} \int_0^\tau \mu(t, \beta_h) dt. \quad (2.18)$$

(2.17) und (2.18) entsprechen jeweils dem n^{-1} -fachen des Zählers und Nenners von $\hat{\lambda}_{0h}(\beta_h)$. Aufgrund des Continuous-Mapping-Theorems (vgl. Billingsley (1969), S. 31, Korollar 1) gilt demnach

$$\hat{\lambda}_{0h}(\beta_h) = \frac{\frac{1}{n} \sum_{i=1}^n N_{hi}(\tau)}{\frac{1}{n} \sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp\left(\hat{\beta}_h^\top \mathbf{X}_i(t)\right) dt} \xrightarrow{\mathcal{P}} \frac{\mathbb{E}[N_{h1}(\tau)]}{\int_0^\tau \mu(t, \beta_h) dt} =: \lambda_{0h}(\beta_h).$$

Der Zähler von $\lambda_{0h}(\beta_h)$ hängt lediglich von $\beta_h^{(0)}$ ab und stellt somit eine Konstante bezüglich β_h dar. Da der Nenner von $\lambda_{0h}(\beta_h)$ eine in β_h stetige Funktion ist, folgt demnach die Behauptung. \square

Für den Nachweis, dass $\lambda_{0h}(\beta_h^{(0)}) = \lambda_{0h}^{(0)}$ ist, bedarf es zunächst einiger Vorüberlegungen in Form der folgenden Lemmas 2.1 bis 2.3.

Lemma 2.1 *Die Annahmen **(A4)** und **(A5)** seien erfüllt. Sei $h \in \{+1, -1\}$ und sei $U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \beta_h^{(0)}, \tau) := \partial/\partial\lambda_{0h} \log L(\lambda_{0h}^{(0)}, \beta_h^{(0)})$ die erste Ableitung der log-Likelihood (2.12) nach λ_{0h} an der Stelle der wahren Parameterwerte. Dann gilt*

$$\frac{1}{n} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \beta_h^{(0)}, \tau) \xrightarrow{\mathcal{P}} 0.$$

Beweis: Sei $s \in [0, \tau]$ und sei $\alpha_{hi}(t) = \lambda_{0h}^{(0)} \exp(\beta_h^{(0)\top} \mathbf{X}_i(t))$. Dann gilt

$$\begin{aligned} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \beta_h^{(0)}, \tau) &= \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) dN_{hi}(t) \\ &\quad - \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \alpha_{hi}(t) dt \\ &= \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) dN_{hi}(t) \end{aligned} \quad (2.19)$$

$$\begin{aligned} &\quad - \sum_{i=1}^n \int_0^s Y_{hi}(t) \left[\frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) \right] \alpha_{hi}(t) dt \\ &= \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) dM_{hi}(t), \end{aligned} \quad (2.20)$$

wobei

$$M_{hi}(t) = N_{hi}(t) - \int_0^t Y_{hi}(u) \alpha_{hi}(u) du.$$

$Y_{hi}(t)\alpha_{hi}(t)$ ist der vorhersagbare Intensitätsprozess von $N_{hi}(t)$. Somit ist $M_{hi}(t)$ nach Andersen et al. (1993), Satz II.4.1, ein lokal quadratisch integrierbares Martingal bezüglich der Filtration $(\mathcal{F}_t^{(n)}, t \in [0, \tau])$. Außerdem gilt wegen **(A4)**

$$0 \leq Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) = Y_{hi}(t) \frac{1}{\lambda_{0h}^{(0)}} < \frac{1}{c_h^l},$$

d.h. der Integrand aus (2.20) ist nach oben und unten beschränkt. Daher ist

$$\int_0^s Y_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) dM_{hi}(t)$$

nach Andersen et al. (1993), Satz II.4.1, ebenfalls ein lokal quadratisch integrierbares Martingal für alle $s \in [0, \tau]$. Somit gilt dies offensichtlich auch für $n^{-1}U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s)$. Der vorhersagbare Variationsprozess von $U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s)$ ist gegeben durch

$$\begin{aligned} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \rangle &= \int_0^s \sum_{i=1}^n Y_{hi}(t) \left[\frac{\partial}{\partial \lambda_{0h}^{(0)}} \log(\alpha_{hi}(t)) \right]^2 \alpha_{hi}(t) dt \\ &= \frac{1}{\lambda_{0h}^{(0)}} \int_0^s \sum_{i=1}^n Y_{hi}(t) \exp(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t)) dt \end{aligned} \quad (2.21)$$

Nun gilt aufgrund von **(A5)** (vgl. Andersen et al. (1993), S.84)

$$\begin{aligned} \frac{1}{n} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \rangle &= \frac{1}{\lambda_{0h}^{(0)}} \int_0^\tau \frac{1}{n} \sum_{i=1}^n Y_{hi}(t) \exp(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t)) dt \\ &\xrightarrow{\mathcal{P}} \frac{1}{\lambda_{0h}^{(0)}} \int_0^\tau \mu_h(t, \boldsymbol{\beta}_h^{(0)}) dt =: \sigma_h(\tau, \boldsymbol{\beta}_h^{(0)}) < \infty. \end{aligned} \quad (2.22)$$

Dies hat wiederum zur Folge, dass

$$\frac{1}{n^2} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \rangle = \left\langle \frac{1}{n} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \right\rangle \xrightarrow{\mathcal{P}} 0. \quad (2.23)$$

Sei nun $\epsilon > 0$ beliebig. Dann folgt mit der Ungleichung von Lenglart (1977) (vgl. Andersen et al. (1993), Satz 2.5.18):

$$P \left(\sup_{s \in [0, \tau]} \left| \frac{1}{n} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \right| > \epsilon \right) \leq \frac{\delta}{\epsilon^2} + P \left(\left\langle \frac{1}{n} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \right\rangle > \delta \right) \forall \delta > 0 \quad (2.24)$$

Der zweite Summand auf der rechten Seite von (2.24) verschwindet wegen (2.23) für $n \rightarrow \infty$. Mit $\delta = \epsilon^3$ ergibt sich also

$$\lim_{n \rightarrow \infty} P \left(\sup_{s \in [0, \tau]} \left| \frac{1}{n} U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \right| > \epsilon \right) \leq \epsilon.$$

Da $\epsilon > 0$ beliebig gewählt wurde, gilt $n^{-1}U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \xrightarrow{\mathcal{P}} 0$ für alle $s \in [0, \tau]$ und somit folgt die Behauptung. \square

Lemma 2.2 Die Annahmen (A4) und (A5) seien erfüllt. Sei $\boldsymbol{\lambda}_0 = (\lambda_{0(+1)}, \lambda_{0(-1)})^\top$ und $\boldsymbol{\lambda}_0^{(o)} = (\lambda_{0(+1)}^{(o)}, \lambda_{0(-1)}^{(o)})^\top$. Dann gilt

$$\frac{1}{n} \frac{\partial^2}{\partial \boldsymbol{\lambda}_0 \partial \boldsymbol{\lambda}_0^\top} \log L(\boldsymbol{\theta}, \tau) \Big|_{(\boldsymbol{\lambda}_0^{(o)\top}, \boldsymbol{\beta}^{(o)\top})} \xrightarrow{\mathcal{P}} - [\sigma_{+1}(\tau, \boldsymbol{\beta}_{+1}^{(o)}), \sigma_{-1}(\tau, \boldsymbol{\beta}_{-1}^{(o)})] \mathbf{I}_2$$

Beweis: Offensichtlich gilt

$$\frac{\partial^2}{\partial \lambda_{0(+1)} \partial \lambda_{0(-1)}} \log L(\boldsymbol{\theta}, \tau) = \frac{\partial^2}{\partial \lambda_{0(-1)} \partial \lambda_{0(+1)}} \log L(\boldsymbol{\theta}, \tau) = 0$$

Sei $s \in [0, \tau]$ und $h \in \{+1, -1\}$. Die zweite Ableitung von $\log L(\boldsymbol{\theta}_h, s)$ nach λ_{0h} an der Stelle der wahren Parameterwerte ist gegeben durch

$$\frac{1}{n} \frac{\partial^2}{(\partial \lambda_{0h})^2} \log L(\boldsymbol{\theta}, s) \Big|_{(\lambda_{0h}^{(o)}, \boldsymbol{\beta}_h^{(o)\top})} = \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h}^{(o)})^2} \log(\alpha_{hi}(t)) dN_{hi}(t) \quad (2.25)$$

$$\begin{aligned} & - \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h})^2} \alpha_{hi}(t) dt \\ & = \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h}^{(o)})^2} \log(\alpha_{hi}(t)) dM_{hi}(t) \quad (2.26) \\ & + \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \alpha_{hi}(t) \frac{\partial}{\partial \lambda_{0h}^{(o)}} \frac{\frac{\partial}{\partial \lambda_{0h}^{(o)}} \alpha_{hi}(t)}{\alpha_{hi}(t)} \alpha_{hi}(t) dt \\ & - \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h}^{(o)})^2} \alpha_{hi}(t) dt. \end{aligned}$$

Der zweite Summand von (2.26) lässt sich umschreiben zu

$$\frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\alpha_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h}^{(o)})^2} \alpha_{hi}(t) - \left(\frac{\partial}{\partial \lambda_{0h}^{(o)}} \alpha_{hi}(t) \right)^2}{\alpha_{hi}(t)} dt.$$

Damit vereinfacht sich (2.26) zu

$$\begin{aligned} \frac{1}{n} \frac{\partial^2}{(\partial \lambda_{0h})^2} \log L(\boldsymbol{\theta}, s) \Big|_{(\lambda_{0h}^{(o)}, \boldsymbol{\beta}_h^{(o)\top})} & = \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\partial^2}{(\partial \lambda_{0h}^{(o)})^2} \log(\alpha_{hi}(t)) dM_{hi}(t) \quad (2.27) \\ & - \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{\left(\frac{\partial}{\partial \lambda_{0h}^{(o)}} \alpha_{hi}(t) \right)^2}{\alpha_{hi}(t)} dt \\ & = \underbrace{-\frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \frac{1}{(\lambda_{0h}^{(o)})^2} dM_{hi}(t)}_{=: A_h(\lambda_{0h}^{(o)}, \boldsymbol{\beta}_h^{(o)}, s)} - \frac{1}{n} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(o)}, \boldsymbol{\beta}_h^{(o)}, s) \rangle. \end{aligned}$$

Für die Integranden von $A_h(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s)$ gilt wegen **(A4)**:

$$-\frac{1}{(c_h^l)^2} < -Y_{hi}(t) \frac{1}{(\lambda_{0h}^{(0)})^2} \leq 0$$

Somit ist $A_h(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s)$ nach Andersen et al. (1993), Satz II.4.1, ein lokal quadratisch integrierbares Martingal mit Variationsprozess

$$\langle A(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \rangle = \frac{1}{n(\lambda_{0h}^{(0)})^3} \int_0^s \frac{1}{n} \sum_{i=1}^n Y_{hi}(t) \exp\left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t)\right) dt,$$

wobei wegen **(A5)** $\langle A_h(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \rangle \xrightarrow{\mathcal{P}} 0$ für $n \rightarrow \infty$. Analog zum Beweis von Lemma 1 folgt demnach mit einer Anwendung der Ungleichung von Lenglar (1977), dass $A_h(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \xrightarrow{\mathcal{P}} 0$. Nach Gleichung (2.22) gilt außerdem, dass $\frac{1}{n} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \rangle \xrightarrow{\mathcal{P}} \sigma_h(\tau, \boldsymbol{\beta}_h^{(0)})$. Mit (2.27) und $s = \tau$ folgt also die Behauptung. \square

Lemma 2.3 *Die Annahmen **(A4)** und **(A5)** seien erfüllt. Dann existiert für $h \in \{+1, -1\}$ eine Konstante $K_h \in (0, \infty)$, so dass*

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{1}{n} \frac{\partial^3}{(\partial \lambda_{0h})^3} \log L(\boldsymbol{\theta}, \tau) \Big|_{\boldsymbol{\beta}_h^{(0)}} \right| < K_h \right) = 1$$

für alle $\lambda_{0h} \in (c_h^l, c_h^u)$. K_h hängt dabei nicht von λ_{0h} ab.

Beweis: Sei $s \in [0, \tau]$ und $h \in \{+1, -1\}$. Dann gilt

$$\begin{aligned} \frac{1}{n} \frac{\partial^3}{(\partial \lambda_{0h})^3} \log L(\boldsymbol{\theta}, s) \Big|_{\boldsymbol{\beta}^{(0)}} &= \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{\partial^3}{(\partial \lambda_{0h})^3} \log \left(\lambda_{0h} \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) \right) dN_{hi}(t) \quad (2.28) \\ &\quad - \frac{1}{n} \sum_{i=1}^n \int_0^s Y_{hi}(t) \underbrace{\frac{\partial^3}{(\partial \lambda_{0h})^3} \left[\lambda_{0h} \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) \right]}_{=0} dt \\ &= \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{1}{\lambda_{0h}^3} dN_{hi}(t). \quad (2.29) \end{aligned}$$

Wegen **(A4)** ist

$$\frac{1}{n} \sum_{i=1}^n \int_0^s \frac{1}{\lambda_{0h}^3} dN_{hi}(t) - \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{\lambda_{0h}^{(0)}}{\lambda_{0h}^3} Y_{hi}(t) \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) dt$$

nach Andersen et al. (1993), Satz II.4.1, ein lokal quadratisch integrierbares Martingal, dessen vorhersagbarer Variationsprozess mit (2.21) gegeben ist durch

$$\frac{1}{n^2} \sum_{i=1}^n \int_0^s \frac{\lambda_{0h}^{(0)}}{\lambda_{0h}^6} Y_{hi}(t) \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) dt = \frac{(\lambda_{0h}^{(0)})^2}{\lambda_{0h}^6 n^2} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, s) \rangle.$$

Sei $\epsilon > 0$ beliebig. Analog zum Beweis von Lemma 2.1 folgt dann mit der Ungleichung von Lengart (1977):

$$\begin{aligned} & P \left(\sup_{s \in [0, \tau]} \left| \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{1}{\lambda_{0h}^3} dN_{hi}(t) - \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{\lambda_{0h}^{(0)}}{\lambda_{0h}^3} Y_{hi}(t) \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) dt \right| > \epsilon \right) \\ & \leq \epsilon + P \left(\frac{(\lambda_{0h}^{(0)})^2}{\lambda_{0h}^6 n^2} \langle U_{\lambda_{0h}}(\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)}, \tau) \rangle > \epsilon^3 \right) \end{aligned} \quad (2.30)$$

Wegen **(A5)** und (2.23) verschwindet der zweite Summand auf der rechten Seite von (2.30) für $n \rightarrow \infty$, d.h. es gilt

$$\frac{1}{n} \sum_{i=1}^n \int_0^s \frac{1}{\lambda_{0h}^3} dN_{hi}(t) - \frac{1}{n} \sum_{i=1}^n \int_0^s \frac{\lambda_{0h}^{(0)}}{\lambda_{0h}^3} Y_{hi}(t) \exp \left(\boldsymbol{\beta}_h^{(0)\top} \mathbf{X}_i(t) \right) dt \xrightarrow{\mathcal{P}} 0.$$

Somit folgt aufgrund von (2.22)

$$\frac{1}{n} \sum_{i=1}^n \int_0^s \frac{1}{\lambda_{0h}^3} dN_{hi}(t) \xrightarrow{\mathcal{P}} \frac{(\lambda_{0h}^{(0)})^2}{\lambda_{0h}^3} \sigma_h(\tau, \boldsymbol{\beta}_h) < \frac{(c_h^u)^2}{(c_h^l)^3} \sigma_h(\tau, \boldsymbol{\beta}_h^{(0)}) =: K_h$$

und damit wegen (2.29) die Behauptung. \square

Satz 2.3 Die Annahmen **(A4)** und **(A5)** seien erfüllt. Dann gilt für $h \in \{-1, +1\}$

$$\hat{\lambda}_{0h}(\boldsymbol{\beta}_h^{(0)}) \xrightarrow{\mathcal{P}} \lambda_{0h}^{(0)}.$$

Beweis: Die Behauptung folgt nach Billingsley (1961), S. 12ff (vgl. auch Andersen et al. (1993), Satz VI.1.1) aus Lemma 2.1 bis 2.3. \square

Satz 2.4 Die Annahmen **(A1)** bis **(A5)** seien erfüllt. Dann gilt für $h \in \{+1, -1\}$

$$\hat{\lambda}_{0h}(\hat{\boldsymbol{\beta}}_h) \xrightarrow{\mathcal{P}} \lambda_{0h}^{(0)},$$

d.h. $\hat{\lambda}_{0h}(\hat{\boldsymbol{\beta}}_h)$ aus (2.14) ist ein konsistenter Schätzer für $\lambda_{0h}^{(0)}$.

Beweis: Wegen **(A1)** und **(A6)** ist

$$\hat{\lambda}_{0h}(\boldsymbol{\beta}_h) = \frac{\sum_{i=1}^n N_{hi}(\tau)}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t) \right) dt}$$

fast sicher differenzierbar bezüglich $\boldsymbol{\beta}_h$ für alle $\boldsymbol{\beta}_h \in \mathbb{R}^d$. Sei \mathcal{B}_h die kompakte Umgebung aus **(A5)** und seien $\boldsymbol{\beta}_h^{(1)}, \boldsymbol{\beta}_h^{(2)} \in \mathcal{B}_h^*$ mit $\mathcal{B}_h^* \supset \mathcal{B}_h$ offen, konvex und beschränkt. Dann gilt nach dem Mittelwertsatz

$$\hat{\lambda}_{0h}(\boldsymbol{\beta}_h^{(1)}) - \hat{\lambda}_{0h}(\boldsymbol{\beta}_h^{(2)}) = \sum_{j=1}^d \frac{\partial \hat{\lambda}_{0h}(\boldsymbol{\beta}_h)}{\partial \beta_{hj}} \Bigg|_{\boldsymbol{\beta}_h^{(3)}} \left(\beta_{hj}^{(1)} - \beta_{hj}^{(2)} \right), \quad (2.31)$$

wobei $\beta_h^{(3)} \in \mathcal{B}_h^*$ ein Punkt auf der Verbindungslinie zwischen $\beta_h^{(1)}$ und $\beta_h^{(2)}$ ist und β_{hj} die j te Komponenten von β_h bezeichnet, $j = 1, \dots, d$. Schreibt man die rechte Seite von (2.31) als Skalarprodukt zweier Vektoren ergibt sich mit Hilfe der Cauchy-Schwarz-Ungleichung

$$\begin{aligned} \left| \hat{\lambda}_{0h}(\beta_h^{(1)}) - \hat{\lambda}_{0h}(\beta_h^{(2)}) \right| &= \left| \left(\frac{\partial \hat{\lambda}_{0h}(\beta_h)}{\partial \beta_h} \Big|_{\beta_h^{(3)}} \right)^\top (\beta_h^{(1)} - \beta_h^{(2)}) \right| \\ &\leq \left\| \frac{\partial \hat{\lambda}_{0h}(\beta_h)}{\partial \beta_h} \Big|_{\beta_h^{(3)}} \right\| \|(\beta_h^{(1)} - \beta_h^{(2)})\| \\ &\leq \sup_{\beta_h^* \in \mathcal{B}_h^*} \left\| \frac{\partial \hat{\lambda}_{0h}(\beta_h)}{\partial \beta_h} \Big|_{\beta_h^*} \right\| \|(\beta_h^{(1)} - \beta_h^{(2)})\|. \end{aligned} \quad (2.32)$$

Die partiellen Ableitungen von $\hat{\lambda}_{0h}$ sind für $j = 1, \dots, d$ gegeben durch

$$\begin{aligned} \frac{\partial \hat{\lambda}_{0h}(\beta_h)}{\partial \beta_{hj}} &= \frac{\sum_{i=1}^n N_{hi}(\tau)}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt} \\ &\quad \times \frac{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) X_{ij}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt}. \end{aligned}$$

Aufgrund der Beschränktheit von $\mathbf{X}_i(t)$, folgt mit dem Gesetz der großen Zahl

$$\frac{1}{n} \sum_{i=1}^n \int_0^\tau Y_{hi}(t) X_{ij}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt \xrightarrow{\mathcal{P}} \mathbb{E} \left[\int_0^\tau Y_{hi}(t) X_{ij}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt \right] < \infty.$$

Mit dem Continuous-Mapping-Theorem ergibt sich dann

$$\frac{\partial \hat{\lambda}_{0h}(\beta_h)}{\partial \beta_{hj}} \xrightarrow{\mathcal{P}} \lambda_{0h}(\beta_h) \cdot \frac{\mathbb{E} \left[\int_0^\tau Y_{hi}(t) X_{ij}(t) \exp(\beta_h^\top \mathbf{X}_i(t)) dt \right]}{\int_0^\tau \mu_h(t, \beta_h) dt} < \infty,$$

d.h. das Supremum (2.32) ist für $n \rightarrow \infty$ fast sicher beschränkt. Da das Supremum (2.32) zudem nicht von $\beta_h^{(1)}$ und $\beta_h^{(2)}$ abhängt, folgt mit Davidson (1994), Satz 21.10, dass $\hat{\lambda}_{0h}(\beta_h)$ gleichgradig stetig auf \mathcal{B}_h^* ist. Aufgrund der Stetigkeit von $\lambda_{0h}(\beta_h)$ (Satz 2.2) ist $\hat{\lambda}_{0h}(\beta_h) - \lambda_{0h}(\beta_h)$ ebenfalls gleichgradig stetig auf \mathcal{B}_h^* und damit auch auf \mathcal{B}_h . Satz 2.2 besagt zudem, dass $\hat{\lambda}_{0h}(\beta_h)$ für alle $\beta_h \in \mathcal{B}_h$ punktweise gegen $\lambda_{0h}(\beta_h)$ konvergiert. Zusammen mit der gleichgradigen Stetigkeit von $\hat{\lambda}_{0h}(\beta_h) - \lambda_{0h}(\beta_h)$ ergibt sich nach Davidson (1994), Satz 21.9, die gleichmäßige Konvergenz, d.h.

$$\sup_{\beta_h \in \mathcal{B}_h} \left| \hat{\lambda}_{0h}(\beta_h) - \lambda_{0h}(\beta_h) \right| \xrightarrow{\mathcal{P}} 0. \quad (2.33)$$

Außerdem gilt nach Satz 2.1, dass $\hat{\beta}_h \xrightarrow{\mathcal{P}} \beta_h^{(0)}$. Zusammen mit (2.33) folgt daraus nach Davidson (1994), Satz 21.6, dass $\hat{\lambda}_{0h}(\hat{\beta}_h)$ in Wahrscheinlichkeit punktweise gegen $\lambda_{0h}(\beta_h^{(0)}) = \lambda_{0h}^{(0)}$ konvergiert. \square

3 Messfehler in den Ratingdaten

In der Praxis ist es nicht realistisch, dass die Bonität eines Schuldners tatsächlich kontinuierlich kontrolliert werden kann. Geht man davon aus, dass die Ratings lediglich in unregelmäßigen Abständen durch die Bank überprüft werden, so werden Ratingwechsel im Mittel verzögert erfasst, wobei diese Verzögerung mal mehr und mal weniger stark ausfallen kann. Dies führt dazu, dass bei der Ermittlung der Ratingwechselzeitpunkte zufällige Messfehler auftreten. In diesem Fall steckt also latente Heterogenität in den Ratingdaten, die sich nicht in Form von Kovariablen berücksichtigen lässt.

3.1 Modellierung von Messfehlern

Messfehler in den Ratingwechselzeitpunkten lassen sich durch einen geeigneten Frailty-Term - also einen latenten stochastischen Prozess - in den Intensitäten berücksichtigen. Seien $\mathbf{V}_h = (V_{hj}, i \in \{1, \dots, n\}, j \in \{1, \dots, N_{hi}(\tau) + 1\})$, $h \in \{+1, -1\}$, zwei unabhängige Familien von positiven u.i.v. Zufallsvariablen mit Erwartungswert 1 und sei

$$W_{hi}(t) = V_{hi(N_{hi}(t^-)+1)}. \quad (3.1)$$

Dann ist ein mögliches Intensitätsmodell für einstufige Upgrades bzw. Downgrades unter Berücksichtigung von Messfehlern gegeben durch

$$\lambda_{hi}^*(t, W_{hi}(t)) = W_{hi}(t)\lambda_{hi}(t). \quad (3.2)$$

Der Prozess $W_{hi}(t)$ ist demnach stückweise konstant mit zufälligem Startwert V_{hi1} und nimmt unmittelbar nach dem j ten Sprung vom Typ h einen neuen zufälligen Wert $V_{hi(j+1)}$ an. Im Modell (3.2) wird also angenommen, dass Messfehler dadurch zustande kommen, dass sich die tatsächliche datenerzeugende Intensität von Sprung zu Sprung verändert und zufällig um die ursprünglich angenommene Intensität $\lambda_{hi}(t)$ schwankt. Nach Kim et al. (2012) verweilen Schuldner in wirtschaftlich schlechten Zeiten tendenziell länger in einer Ratingklasse. Dies legt die Vermutung nahe, dass die Bonität der Schuldner in wirtschaftlich schlechten Zeiten in größeren Abständen überprüft wird, so dass Messfehler bei den Downgrades im Mittel größer ausfallen könnten. Daher ergibt es Sinn, bei den Familien \mathbf{V}_{+1} und \mathbf{V}_{-1} im Allgemeinen von einer unterschiedlichen Verteilung auszugehen.

Aus der Intensität (3.2) ergibt sich nun - wie in Kapitel 2.2 beschrieben - eine Likelihood, welche allerdings von den Prozessen $W_{hi}(t)$ abhängt. Diese sind - im Gegensatz zu den Prozessen $Y_{hi}(t)$ und $\mathbf{X}_i(t)$ - jedoch nicht beobachtbar, so dass die Likelihood nicht ohne weiteres für die Parameterinferenz verwendet werden kann. Unter der Annahme, dass die Zensierungszeiten der Schuldner unabhängig von den Migrationszeitpunkten und dem Prozess $W_{hi}(t)$ sind, lässt sich jedoch zeigen, dass die Likelihood - bis auf einen bezüglich der Modellparameter irrelevanten Faktor - der bedingten Dichte der Ereigniszeiten bei gegebenem Messfehlerprozess entspricht. Nach den Regeln der Wahrscheinlichkeitsrechnung ergibt sich daraus die Randdichte der Ereigniszeiten, indem der Erwartungswert der bedingten Dichte bezüglich der Messfehlerprozesse $W_{hi}(t)$ gebildet wird. Diese Randdichte lässt sich als marginale Likelihood für die Parameterinferenz verwenden. Eine detaillierte Herleitung dieser marginalen Likelihood für den Spezialfall eines zeitkonstanten gamma-verteilten Frailty-Terms liefern Nielsen et al. (1992) mit Hilfe der Zählprozessentheorie.

Bei der Bestimmung des Erwartungswerts der Likelihood stellt sich die Frage nach der Verteilung der Zufallsvariablen V_{hij} , die den Prozessen $W_{hi}(t)$ zugrunde liegen. Eine nahe liegende Möglichkeit besteht darin, dieser ein parametrisches Modell zu unterstellen und die entsprechenden Parameter zusammen mit λ_{0h} und β_h zu schätzen. Allerdings hängen die Schätzer für λ_{0h} und β_h nach Heckman und Singer (1984b) und Heckman und Singer (1984a) sehr stark von der parametrischen Form dieser Verteilung ab. Heckman und Singer (1984b) schlagen ein nichtparametrisches Verfahren als Alternative vor. Dabei wird angenommen, dass die Verteilung der Frailty-Terme diskret ist. Auf diese Weise lässt sich der Erwartungswert als Summe darstellen. Die Anzahl der Trägerpunkte wird zunächst auf 2 gesetzt. Dann werden die Trägerpunkte, die zugehörige Wahrscheinlichkeitsmasse und die Modellparameter durch Maximierung der erwarteten Likelihood geschätzt. Anschließend wird die Anzahl der Trägerpunkte um einen erhöht und die Schätzung erneut durchgeführt. Dies wird solange wiederholt, bis sich der Wert der erwarteten Likelihood nicht mehr erhöht. Dieses Verfahren ist recht aufwendig, da die Anzahl der zu schätzenden Parameter sehr hoch werden kann. Dies erschwert die numerische Bestimmung des Maximums. Die Wahl der Startwerte des numerischen Verfahrens gewinnt mit jedem Iterationsschritt an Bedeutung. Die späteren Iterationsschritte sollten also mehrfach mit verschiedenen Startwerten durchgeführt werden, um sicherzugehen, ein globales Maximum zu finden. Eine detaillierte Beschreibung des Verfahrens und dessen Eigenschaften liefern Trussel und Richards (1985).

Zusammenfassend lässt sich also sagen, dass die Berücksichtigung von Messfehlern recht aufwendig ist und einer gewissen Willkür unterliegt, da die Ergebnisse stark von der Wahl des Verfahrens zur Bestimmung der marginalen Likelihood abhängen können. Um deutlich zu machen, dass die Berücksichtigung von Messfehlern trotz der beschriebenen Komplika-

tionen von entscheidender Bedeutung ist, werden im folgenden Abschnitt die Auswirkungen der Vernachlässigung aufgezeigt.

3.2 Folgen der Vernachlässigung

Im Folgenden werden die Auswirkung der Vernachlässigung von zeitkonstanten Kovariablen oder Messfehlern in den Ratingwechselzeitpunkten auf die Parameterschätzer anhand einer Simulationsstudie überprüft.¹ Ausgangspunkt für diese Studie ist ein Ratingsystem mit 21 Ratingklassen, in dessen Rahmen mit Hilfe der Intensitäten $\lambda_{hi}^*(t, W_{hi}(t))$ Ratinghistorien von $n = 250$ Schuldnern simuliert werden. Das Startrating wird dabei jeweils zufällig vergeben. Die Beobachtungsdauer eines Schuldners wird durch eine zensierte $\text{Exp}(0.4)$ -verteilte Zufallsvariable erzeugt. Somit beträgt die Länge der Ratinghistorien im Mittel etwas weniger als zweieinhalb Jahre und im Höchstfall 10 Jahre. Die Intensitäten hängen jeweils von einem dreidimensionalen Kovariablenvektor $\mathbf{X}_i(t) = (X_{1i}, X_{2i}, X_{3i}(t))^\top$ ab. Dabei stammt X_{1i} aus einer $N(0, 0.75)$ -Verteilung, während X_{2i} als $\text{Bin}(1, 0.5)$ -verteilt angenommen wird. Die zeitabhängige Kovariable $X_{3i}(t)$ wird durch einen stationären $\text{AR}(2)$ -Prozess simuliert. Den Familien \mathbf{V}_h , welche die Basis für den Messfehlerprozess darstellen, wird eine Gamma-Verteilung mit Erwartungswert 1 unterstellt. Die Schätzung der acht Intensitätsparameter $\lambda_{0(+1)}$, $\beta_{(+1)1}$, $\beta_{(+1)2}$, $\beta_{(+1)3}$, $\lambda_{0(-1)}$, $\beta_{1(-1)}$, $\beta_{1(-2)}$ und $\beta_{1(-3)}$ erfolgt dann an Hand der erzeugten Daten mit den in Kapitel 2.2 beschriebenen Methoden. Von Interesse ist dabei das Verhalten der Schätzer mit steigender Varianz der den Messfehlern unterstellten Gamma-Verteilung. Die Schätzer für die Upgrade- und Downgradeparameter reagieren gleichermaßen auf eine Erhöhung der Messfehlervarianz, da beide Intensitäten dieselbe Struktur aufweisen. Daher werden aus Gründen der Übersichtlichkeit im Folgenden lediglich die Downgradeschätzer betrachtet.

Abbildung 3.1 zeigt das Verhalten der Downgradeschätzer bei der Missachtung von Messfehlern. Es ist zu beobachten, dass die Varianz der Parameterschätzer mit steigender Messfehlervolatilität zunimmt. Außerdem fällt auf, dass die Kovariableneinflüsse $\beta_{(+1)1}$, $\beta_{(+1)2}$ und $\beta_{(+1)3}$ unterschätzt werden, wobei der Grad der Verzerrung vom Kovariablentyp abzuhängen scheint. Am größten fällt der Fehler bei dem Einfluss der stetig-skalierten zeitkonstanten Kovariable aus. Hier ergibt sich bei einer Messfehlervarianz von 2 ein Schätzfehler von 44%. Bei der dichotomen Kovariable beträgt dieser lediglich rund 30%, wohingegen er bei der zeitabhängigen Kovariable mit 23% am geringsten ausfällt. $\lambda_{0(-1)}$ wird mit zunehmender Messfehlervarianz ebenfalls immer stärker unterschätzt. Bei einer Messfehlervarianz von 2 ergibt sich ein Schätzfehler von 37%. Die Streuung von $\hat{\lambda}_{0(-1)}$ bleibt

¹Alle im Verlauf dieser Arbeit vorgestellten Simulationsstudien und empirische Auswertungen wurden mit der Statistiksoftware R durchgeführt.

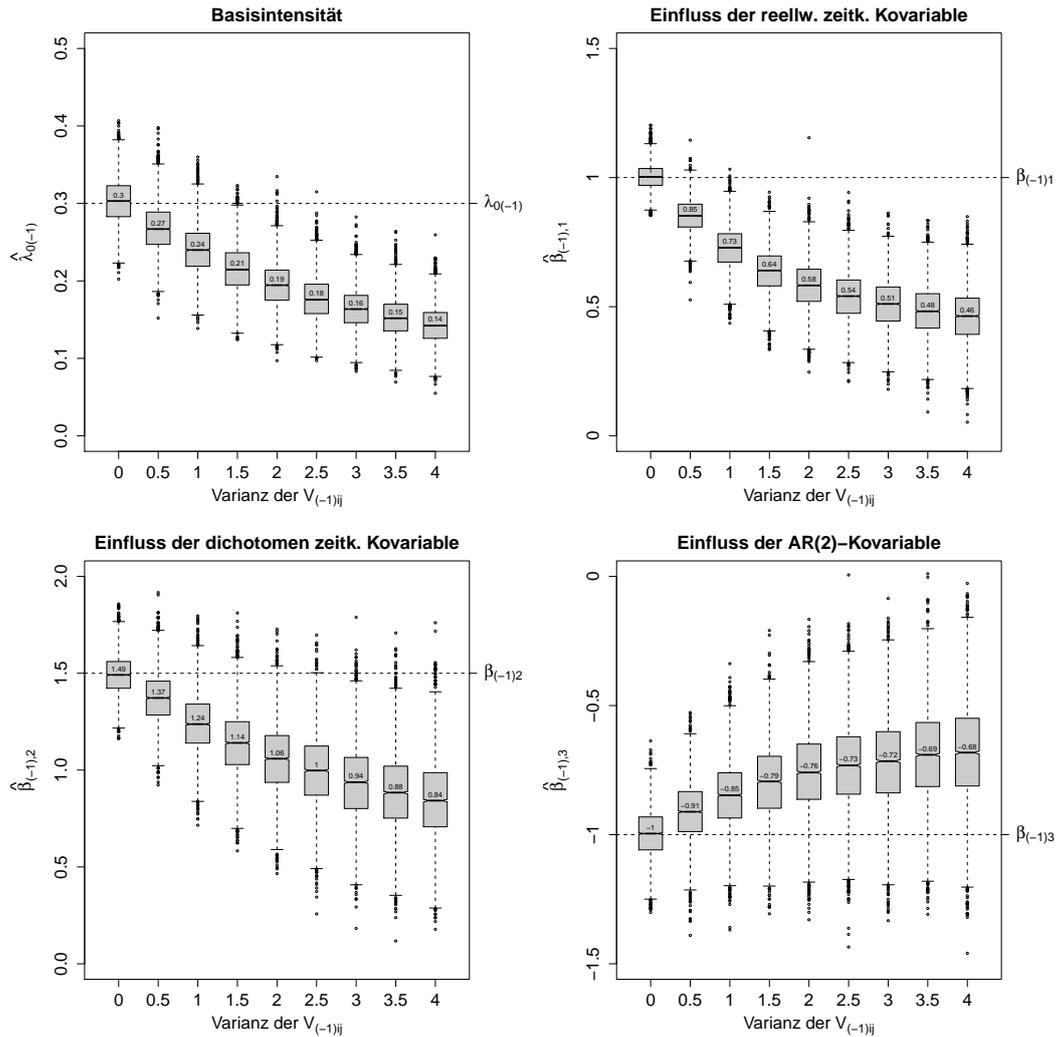


Abbildung 3.1: Lage und Streuung der Schätzer für die Parameter der Downgrade-Intensität $\lambda_{-1i}(t)$ bei Vernachlässigung von Messfehlern ($n = 250$)

allerdings nahezu unverändert.

Insgesamt lässt sich also sagen, dass durch die Vernachlässigung von Messfehlern sowohl die Intensitäten, als auch die durch die Kovariablen verursachten Schwankungen in den Intensitäten unterschätzt werden. Dies hat zur Folge, dass auch die Ratingwechselwahrscheinlichkeiten zu niedrig eingeschätzt werden. Aufgrund der Tatsache, dass die Wechselwahrscheinlichkeiten einen Inputfaktor bei der Bestimmung der Verlustverteilung darstellen, führt die Missachtung von Messfehlern zu einer Fehleinschätzung des Kreditrisikos seitens der Bank und kann somit schwerwiegende Konsequenzen haben. Da die Berücksichtigung von Messungenauigkeiten jedoch mit einem recht großen Aufwand ein-

hergeht, ist es sinnvoll, zunächst mit Hilfe eines statistischen Tests zu entscheiden, ob die vorliegenden Daten tatsächlich Messfehler enthalten. Ein solcher Test wird im folgenden Abschnitt entwickelt.

4 Score-Test auf Messfehler

Der im Folgenden hergeleitete Test beruht auf einer Idee von Lancaster (1990), Kapitel 11.2.2. Es soll entschieden werden, ob sich die Varianz der dem Messfehlerprozess zugrunde liegenden Zufallsvariablen V_{hij} signifikant von 0 unterscheidet. Ist dies nicht der Fall, gilt $V_{hij} = E(V_{hij}) = 1 \forall hij$ und es liegen keine Messfehler vor. Als Grundlage für einen solchen Test bietet sich die marginale Likelihood der Ereigniszeiten, d.h. der Erwartungswert der Likelihood L bezüglich der V_{hij} an. Dieser Erwartungswert enthält die Varianz der Messfehler als Parameter und kann somit zum Testen verwendet werden. Die Bestimmung dieser marginalen Likelihood bedarf allerdings einer Verteilungsannahme an die V_{hij} , welche die Testentscheidung beeinflussen könnte. Diesem Problem lässt sich aus dem Weg gehen, indem der Erwartungswert von L lediglich geeignet approximiert wird, so dass keine Verteilungsannahme nötig ist. Lancaster (1990) hat einen solchen Test für unzensierte Verweildauern konstruiert. Jaggia (1997) hat diese Idee schließlich auf zensierte Verweildauerdaten erweitert und dessen Eigenschaften für den Fall eines festen Zensierungszeitpunktes analysiert. Ein solcher Test soll nun auch für das in Kapitel 2 beschriebene Ratingmodell bestimmt werden.

4.1 Herleitung des Tests

Als Basis für diesen Test dient die Likelihood, welche aus der Messfehler-Intensität (3.2) hervorgeht. Seien $\mathbf{V}_{hi} = (V_{hij}, j \in \{1, \dots, N_{hi}(\tau) + 1\}) \subset \mathbf{V}_h$ die Messfehlerterme, die dem Prozess $W_{hi}(t)$ zu Grunde liegen. Der i te Faktor der Likelihood des Modells (3.2) ergibt sich analog zu (2.9) und ist gegeben durch

$$\begin{aligned} L_i^*(\boldsymbol{\theta}, \mathbf{V}_{+1i}, \mathbf{V}_{-1i}) &= \prod_{h \in \{-1, +1\}} L_{hi}^*(\boldsymbol{\theta}_h, \mathbf{V}_{hi}) \\ &= \prod_{h \in \{-1, +1\}} \prod_{j=1}^{N_{hi}(\tau)} \lambda_{hi}^*(T_{hij}, W_{hi}(T_{hij})) \\ &\quad \times \exp\left(-\int_0^\tau \lambda_{hi}^*(t, W_{hi}(t)) dt\right). \end{aligned} \quad (4.1)$$

Setze $T_{hi0} := 0$. Dann lässt sich L_{hi}^* schreiben als

$$\begin{aligned} L_{hi}^*(\boldsymbol{\theta}_h, \mathbf{V}_{hi}) &= \prod_{j=1}^{N_{hi}(\tau)} \left[V_{hij} \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij})\right) \right. \\ &\quad \times \exp\left(-V_{hij} \int_{T_{hi(j-1)}}^{T_{hij}} Y_{hi}(t) \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t)\right) dt\right) \\ &\quad \left. \times \exp\left(-V_{hi(N_{hi}(\tau)+1)} \int_{T_{hiN_{hi}(\tau)}}^{\tau} Y_{hi}(t) \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t)\right) dt\right) \right]. \end{aligned} \quad (4.2)$$

L_{hi}^* zerfällt also in $N_{hi}(\tau) + 1$ Faktoren, welche im Folgenden mit $L_{hi1}^*, \dots, L_{hi(N_{hi}(\tau)+1)}^*$ bezeichnet werden. Für $j \in \{1, \dots, N_{hi}(\tau)\}$ entspricht L_{hij}^* der bedingten Dichte einer Ereigniszeit U gegeben $U > T_{hi(j-1)}$, wobei die Verteilungsfunktion von U gegeben ist durch $F(u) = 1 - \exp(-\int_0^u \lambda_{hi}(t; \mathbf{X}_i(t), V_{hij}) dt)$. Der letzte Faktor $L_{hi(N_{hi}(\tau)+1)}^*$ unterscheidet sich in seiner Gestalt von den ersten $N_{hi}(\tau)$ Faktoren und stellt die bedingte Überlebensfunktion von U gegeben $U > T_{hiN_{hi}(\tau)}$ dar. Somit lässt sich L_{hi}^* alternativ als eine Likelihood von $N_{hi}(\tau) + 1$ trunkierten Realisation der Ereigniszeit U auffassen, von denen die letzte Beobachtung zensiert ist.

Im ersten Schritt wird L_{hij}^* nun für $j = 1, \dots, N_{hi}(\tau) + 1$ jeweils durch ein Taylorpolynom zweiten Grades um $E(V_{hij}) = 1$ ersetzt. Dies entspricht einer vernünftigen Approximation von L_{hij}^* , falls die Messfehlerterme V_{hij} nur eine geringe Variabilität aufweisen und somit Werte in der Nähe von 1 annehmen. Für die Taylorentwicklung gilt es, zunächst die ersten und zweiten Ableitungen von L_{hij}^* an der Stelle $V_{hij} = 1$ zu bestimmen. Sei

$$z_{hij}(\boldsymbol{\theta}_h) := \int_{T_{hi(j-1)}}^{T_{hij}} Y_{hi}(t) \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t)\right) dt \quad (4.3)$$

für $j = 1, \dots, N_{hi}(\tau) + 1$, wobei $T_{hi(N_{hi}(\tau)+1)} := \tau$. Dann gilt für $j = 1, \dots, N_{hi}(\tau)$

$$\left. \frac{\partial L_{hij}^*}{\partial V_{hij}} \right|_{V_{hij}=1} = -\lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij})\right) \exp(-z_{hij}(\boldsymbol{\theta}_h)) [z_{hij}(\boldsymbol{\theta}_h) - 1], \quad (4.4)$$

$$\left. \frac{\partial^2 L_{hij}^*}{(\partial V_{hij})^2} \right|_{V_{hij}=1} = \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij})\right) \exp(-z_{hij}(\boldsymbol{\theta}_h)) [z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)]. \quad (4.5)$$

Eine Taylorentwicklung zweiten Grades liefert

$$\begin{aligned} L_{hij}^*(\boldsymbol{\theta}_h, V_{hij}) &\approx \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij})\right) \exp(-z_{hij}(\boldsymbol{\theta}_h)) \\ &\quad \times \left(1 - (V_{hij} - 1)(z_{hij}(\boldsymbol{\theta}_h) - 1) + \frac{(V_{hij} - 1)^2}{2} (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h))\right) \\ &=: \tilde{L}_{hij}^*(\boldsymbol{\theta}_h, V_{hij}). \end{aligned} \quad (4.6)$$

Die ersten und zweiten Ableitungen der $L_{hi(N_{hi}(\tau)+1)}^*$ an der Stelle $V_{hi(N_{hi}(\tau)+1)} = 1$ sind gegeben durch

$$\left. \frac{\partial L_{hij}^*}{\partial V_{hi(N_{hi}(\tau)+1)}} \right|_{V_{hi(N_{hi}(\tau)+1)}=1} = -\exp(-z_{hij}(\boldsymbol{\theta}_h)) z_{hij}(\boldsymbol{\theta}_h), \quad (4.7)$$

$$\left. \frac{\partial^2 L_{hij}^*}{(\partial V_{hi(N_{hi}(\tau)+1)})^2} \right|_{V_{hi(N_{hi}(\tau)+1)}=1} = \exp(-z_{hij}(\boldsymbol{\theta}_h)) z_{hij}^2(\boldsymbol{\theta}_h). \quad (4.8)$$

Damit folgt

$$\begin{aligned} L_{hi(N_{hi}(\tau)+1)}^*(\boldsymbol{\theta}_h, V_{hi(N_{hi}(\tau)+1)}) &\approx \exp(-z_{hi(N_{hi}(\tau)+1)}(\boldsymbol{\theta}_h)) \\ &\times \left(1 - (V_{hi(N_{hi}(\tau)+1)} - 1) z_{hi(N_{hi}(\tau)+1)}(\boldsymbol{\theta}_h) \right. \\ &\quad \left. + \frac{(V_{hi(N_{hi}(\tau)+1)} - 1)^2}{2} z_{hi(N_{hi}(\tau)+1)}^2(\boldsymbol{\theta}_h) \right) \\ &=: \tilde{L}_{hi(N_{hi}(\tau)+1)}^*(\boldsymbol{\theta}_h, V_{hi(N_{hi}(\tau)+1)}). \end{aligned} \quad (4.9)$$

Bei geringer Messfehlervarianz nimmt V_{hij} für alle $j \in \{1, \dots, N_{hi}(\tau) + 1\}$, $h \in \{+1, -1\}$ Werte in der Nähe von 1 an und es gilt

$$L_i^*(\boldsymbol{\theta}, \mathbf{V}_{+1i}, \mathbf{V}_{-1i}) \approx \prod_{h,j} \tilde{L}_{hij}^*(\boldsymbol{\theta}_h, V_{hij}) =: \tilde{L}_i^*(\boldsymbol{\theta}, \mathbf{V}_{+1i}^*, \mathbf{V}_{-1i}^*). \quad (4.10)$$

Da die Messfehlerterme als unabhängig angenommen werden, ist der Erwartungswert von \tilde{L}_i^* gegeben durch

$$\bar{L}_i(\boldsymbol{\theta}, \boldsymbol{\eta}) := \mathbb{E}_{\mathbf{V}} \left[\tilde{L}_i^*(\boldsymbol{\theta}, \mathbf{V}_{+1i}, \mathbf{V}_{-1i}) \right] = \prod_{h \in \{+, -\}} \prod_{j=1}^{N_{hi}(\tau)+1} \mathbb{E}_{V_{hij}} \left[\tilde{L}_{hij}^*(\boldsymbol{\theta}_h, V_{hij}) \right], \quad (4.11)$$

wobei $\boldsymbol{\eta}$ den Parametervektor der Messfehler-Verteilung bezeichnet. Da \tilde{L}_{hij} nach (4.6) und (4.9) linear in V_{hij} und V_{hij}^2 ist, lässt sich der Erwartungswert leicht ohne eine Verteilungsannahme an die V_{hij} bestimmen. Für $j = 1, \dots, N_{hi}(\tau)$ gilt

$$\begin{aligned} \mathbb{E}_{V_{hij}} \left[\tilde{L}_{hij}^*(\boldsymbol{\theta}_h, V_{hij}) \right] &= \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij})\right) \exp(-z_{hij}(\boldsymbol{\theta}_h)) \\ &\times \left(1 - \mathbb{E}_{v_{hi}} [V_{hij} - 1] (z_{hij}(\boldsymbol{\theta}_h) - 1) \right. \\ &\quad \left. + \frac{\mathbb{E}_{v_{hi}} [(V_{hij} - 1)^2]}{2} (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) \right). \end{aligned} \quad (4.12)$$

Mit $\mathbb{E}[V_{hij}] = 1$ und

$$\eta_h := \frac{\text{Var}(V_{hij})}{2} = \frac{\mathbb{E}[(V_{hij} - 1)^2]}{2} \quad (4.13)$$

ergibt sich

$$\begin{aligned} E_{V_{hij}} \left[\tilde{L}_{hij}^*(\boldsymbol{\theta}_h, V_{hij}) \right] &= \lambda_{0h} \exp \left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij}) \right) \exp(-z_{hij}(\boldsymbol{\theta}_h)) \\ &\times \left[1 + \eta_h (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) \right]. \end{aligned} \quad (4.14)$$

Analog folgt

$$\begin{aligned} E_{V_{hi(N_{hi}(\tau)+1)}} \left[\tilde{L}_{hi(N_{hi}(\tau)+1)}^*(\boldsymbol{\theta}_h, V_{hi(N_{hi}(\tau)+1)}) \right] &= \exp(-z_{hi(N_{hi}(\tau)+1)}(\boldsymbol{\theta}_h)) \\ &\times \left[1 + \eta_h z_{hi(N_{hi}(\tau)+1)}^2(\boldsymbol{\theta}_h) \right]. \end{aligned} \quad (4.15)$$

Somit gilt mit (4.14) und (4.15) und aufgrund der Tatsache, dass die n Schuldner als unabhängig angenommen werden

$$\begin{aligned} \bar{L}(\boldsymbol{\theta}, \boldsymbol{\eta}) &:= \prod_{i=1}^n \bar{L}_i(\boldsymbol{\theta}, \boldsymbol{\eta}) \\ &= L(\boldsymbol{\theta}) \prod_{i=1}^n \prod_{h \in \{+1, -1\}} \left[1 + \eta_h z_{hi(N_{hi}(\tau)+1)}^2(\boldsymbol{\theta}_h) \right] \\ &\quad \times \prod_{i=1}^n \prod_{h \in \{+1, -1\}} \prod_{j=1}^{N_{hi}(\tau)} \left[1 + \eta_h (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) \right]. \end{aligned} \quad (4.16)$$

Dabei bezeichnet $L(\boldsymbol{\theta})$ die Likelihood (2.11) aus dem Modell ohne Messfehlerterm.

Man beachte, dass \bar{L} für $\eta_h > 1$ negativ werden kann. Daher ist diese Likelihood nicht zur Parameterschätzung verwendbar. Sie bietet sich jedoch als Grundlage für Tests auf Messfehler an, da sie die Varianz der Messfehlerterme V_{hij} als Parameter enthält und die tatsächliche erwartete Likelihood bei geringer Varianz der V_{hij} gut approximiert. Die zu testende Nullhypothese lautet „ $H_0: \eta_h = 0 \forall h \in \{+1, -1\}$ “. Da $\eta_h = \text{Var}(V_{hij})/2$ gilt, ist dies gleichbedeutend mit „ $V_{hij} = 1 \forall hij$ “, d.h. die Daten enthalten keine Messfehler. Da eine Varianz keinen negativen Wert annehmen kann, liegt der Wert von $\boldsymbol{\eta}$ unter der Nullhypothese am Rand des zulässigen Parameterraums. Dies hat nach Godfrey (1990) einige unschöne Eigenschaften des ML-Schätzers für $\boldsymbol{\eta}$ zur Folge. Da der wahre Wert des Parameters nicht unterschätzt werden kann, ist der ML-Schätzer zum Beispiel nicht asymptotisch normalverteilt ist. Ein Likelihood-Quotienten- oder Wald-Test ist in diesem Fall nicht ohne weiteres durchführbar. Es bietet sich jedoch ein Score-Test an, da dazu lediglich der Parameterschätzer unter der Restriktion der Nullhypothese bestimmt werden muss. Da $\boldsymbol{\eta}$ durch H_0 bereits eindeutig festgelegt ist, muss also lediglich $\boldsymbol{\theta}$ geschätzt werden. Ein Score-Test basiert auf der Idee, dass der Score-Vektor - gegeben durch die erste Ableitung der Log-Likelihood - an der Stelle der wahren Parameterwerte in Erwartung den Wert 0 annimmt. Sofern die Ableitung der Log-Likelihood auch am Rand des Parameterbereichs

existiert und im Erwartungswert verschwindet, lässt sich der Test auch in diesem Fall durchführen. Der Nachweis dieser Eigenschaften erfolgt in Kapitel 4.2.

Sei $\boldsymbol{\theta}^* := (\boldsymbol{\theta}_{+1}, \eta_{+1}, \boldsymbol{\theta}_{-1}, \eta_{-1})$. Grundlage für die Teststatistik des Score-Tests für „ $H_0 : \eta_h = 0 \forall h \in \{+1, -1\}$ “ ist nach Rao (2005) die Score-Funktion von $\boldsymbol{\eta} := (\eta_{+1}, \eta_{-1})$ an der Stelle des Schätzers für $\boldsymbol{\theta}^*$ unter H_0 , d.h.

$$\bar{U}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\theta}}) := \frac{\partial}{\partial \boldsymbol{\eta}} \log \bar{L}(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1,0}^{\top}, \hat{\boldsymbol{\theta}}_{-1,0}^{\top})^{\top}}, \quad (4.17)$$

wobei $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_{+1}^{\top}, \hat{\boldsymbol{\theta}}_{-1}^{\top}) = (\lambda_{0(+1)}, \boldsymbol{\beta}_{+1}^{\top}, \lambda_{0(-1)}, \boldsymbol{\beta}_{-1}^{\top})^{\top}$ die Schätzer aus Kapitel 2.2 bezeichnen. Sei

$$\boldsymbol{\mathcal{I}}(\boldsymbol{\theta}) = -\text{E} \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}^* (\partial \boldsymbol{\theta}^*)^{\top}} \log \bar{L}_1(\boldsymbol{\theta}^*) \right)_{(\boldsymbol{\theta}_{+1,0}^{\top}, \boldsymbol{\theta}_{-1,0}^{\top})^{\top}} \right] \quad (4.18)$$

die Fisher-Informationsmatrix von \bar{L}_i an der Stelle $(\boldsymbol{\theta}_{+1}^{\top}, 0, \boldsymbol{\theta}_{-1}^{\top}, 0)^{\top}$. Ein Schätzer für $\boldsymbol{\mathcal{I}}(\boldsymbol{\theta})$ ist gegeben durch

$$\hat{\boldsymbol{\mathcal{I}}}(\hat{\boldsymbol{\theta}}) = -\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^* (\partial \boldsymbol{\theta}^*)^{\top}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1,0}^{\top}, \hat{\boldsymbol{\theta}}_{-1,0}^{\top})^{\top}}. \quad (4.19)$$

Sei weiter $\boldsymbol{\lambda}_0 = (\lambda_{0(+1)}, \lambda_{0(-1)})^{\top}$ und

$$\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\theta}}) := \hat{\boldsymbol{\mathcal{I}}}_{-\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0}(\hat{\boldsymbol{\theta}}) - \hat{\boldsymbol{\mathcal{I}}}_{-\boldsymbol{\lambda}_0 \boldsymbol{\lambda}_0}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{\mathcal{I}}}_{\boldsymbol{\lambda}_0 \boldsymbol{\lambda}_0}^{-1}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{\mathcal{I}}}_{\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0}(\hat{\boldsymbol{\theta}}), \quad (4.20)$$

wobei die Ausdrücke $\hat{\boldsymbol{\mathcal{I}}}_{\pm \boldsymbol{\lambda}_0 \pm \boldsymbol{\lambda}_0}$ Partitionen von $\hat{\boldsymbol{\mathcal{I}}}$ bezeichnen. Der Index $\boldsymbol{\lambda}_0$ bedeutet dabei, dass ausschließlich die zu den Parametern $\boldsymbol{\lambda}_0$ gehörenden Zeilen bzw. Spalten von $\hat{\boldsymbol{\mathcal{I}}}$ betrachtet werden, während der Index $-\boldsymbol{\lambda}_0$ eine Eliminierung dieser Zeilen bzw. Spalten darstellt. Seien $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}\boldsymbol{\eta}} \in \mathbb{R}^{2 \times 2}$ die Zeilen und Spalten von $\hat{\boldsymbol{\Sigma}}$, die in $\hat{\boldsymbol{\mathcal{I}}}_{-\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_0}$ zu $\boldsymbol{\eta}$ gehören. Dann gilt unter der Nullhypothese und passenden Regularitätsbedingungen, dass

$$\frac{1}{n} \bar{U}_{\boldsymbol{\eta}}^{\top}(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1}(\hat{\boldsymbol{\theta}}) \bar{U}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\theta}}) \stackrel{\mathcal{D}}{\sim} \chi_2^2. \quad (4.21)$$

Der Nachweis erfolgt in Kapitel 4.2. H_0 kann also zum Niveau α verworfen werden, falls (4.21) größer ist als das $(1 - \alpha)$ -Quantil der χ^2 -Verteilung mit 2 Freiheitsgraden. Die einzelnen Komponenten der Teststatistik (4.21) werden im Folgenden bestimmt.

Die logarithmierte Likelihood ist gegeben durch

$$\begin{aligned} \log \bar{L}(\boldsymbol{\theta}^*) = & \sum_{i=1}^n \left(\log L_i(\boldsymbol{\theta}) + \sum_{h \in \{+1, -1\}} \log \left[1 + \eta_h z_{hi}^2(N_{hi}(\tau)+1)(\boldsymbol{\theta}_h) \right] \right. \\ & \left. + \sum_{h \in \{+1, -1\}} \sum_{j=1}^{N_{hi}(\tau)} \log \left[1 + \eta_h (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) \right] \right) \end{aligned} \quad (4.22)$$

mit

$$\log L_i(\boldsymbol{\theta}) = \sum_{h \in \{+1, -1\}} \left(\sum_{j=1}^{N_{hi}(\tau)} \left[\log(\lambda_{0h}) + \boldsymbol{\beta}_h^\top \mathbf{X}_i(T_{hij}) \right] - \sum_{j=1}^{N_{hi}(\tau)+1} z_{hij}(\boldsymbol{\theta}_h) \right). \quad (4.23)$$

Für die Score-Funktion von $\boldsymbol{\eta}$ sind lediglich die letzten beiden Summanden aus (4.22) relevant. Es gilt

$$\bar{U}_{\eta_h}(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^n \left(\sum_{j=1}^{N_{hi}(\tau)} \left(z_{hij}^2(\hat{\boldsymbol{\theta}}_h) - 2z_{hij}(\hat{\boldsymbol{\theta}}_h) \right) + z_{hi(N_{hi}(\tau)+1)}^2(\hat{\boldsymbol{\theta}}_h) \right). \quad (4.24)$$

Um (4.19) zu bestimmen, werden die zweiten Ableitungen von $\log \bar{L}_i$ nach $\boldsymbol{\theta}^*$ an der Stelle $(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top$ benötigt. Da die Likelihood in einen Upgrade- und einen Downgradefaktor zerfällt, verschwinden die bezüglich h gemischten Ableitungen. Die von 0 verschiedenen Komponenten von (4.19) ergeben sich aus

$$\begin{aligned} \frac{\partial^2}{(\partial \eta_h)^2} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= - \sum_{j=1}^{N_{hi}(\tau)} \left(z_{hij}^2(\hat{\boldsymbol{\theta}}_h) - 2z_{hij}(\hat{\boldsymbol{\theta}}_h) \right)^2 - z_{hi(N_{hi}(\tau)+1)}^4(\hat{\boldsymbol{\theta}}_h) \\ \frac{\partial^2}{(\partial \lambda_{0h})^2} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= - \frac{N_{hi}(\tau)}{\hat{\lambda}_{0h}^2} \\ \frac{\partial^2}{\partial \beta_{hl} \partial \beta_{hl'}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= - \hat{\lambda}_{0h} \int_0^\tau Y_{hi}(t) X_{il}(t) X_{il'}(t) \exp\left(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt \\ \frac{\partial^2}{\partial \eta_h \partial \lambda_{0h}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= \sum_{j=1}^{N_{hi}(\tau)} \left(\frac{2}{\hat{\lambda}_{0h}} \left(z_{hij}^2(\hat{\boldsymbol{\theta}}_h) - z_{hij}(\hat{\boldsymbol{\theta}}_h) \right) \right) \\ &\quad + \frac{2}{\hat{\lambda}_{0h}} z_{hi(N_{hi}(\tau)+1)}^2(\hat{\boldsymbol{\theta}}_h) \\ \frac{\partial^2}{\partial \eta_h \partial \beta_{hl}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= \sum_{j=1}^{N_{hi}(\tau)} \left(2 \left(z_{hij}(\hat{\boldsymbol{\theta}}_h) - 1 \right) \frac{\partial}{\partial \beta_{hl}} z_{hij}(\boldsymbol{\theta}_h) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} \right) \\ &\quad + 2z_{hi(N_{hi}(\tau)+1)}(\hat{\boldsymbol{\theta}}_h) \\ &\quad \times \frac{\partial}{\partial \beta_{hl}} z_{hi(N_{hi}(\tau)+1)}(\boldsymbol{\theta}_h) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} \\ \frac{\partial^2}{\partial \lambda_{0h} \partial \beta_{hl}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} &= - \int_0^\tau Y_{hi}(t) X_{il}(t) \exp\left(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt, \end{aligned} \quad (4.25)$$

wobei

$$\frac{\partial}{\partial \beta_{hl}} z_{hij}(\boldsymbol{\theta}_h) \Big|_{(\hat{\boldsymbol{\theta}}_{+1}^\top, 0, \hat{\boldsymbol{\theta}}_{-1}^\top, 0)^\top} = \hat{\lambda}_{0h} \int_{T_{hi(j-1)}}^{T_{hij}} Y_{hi}(t) X_{il}(t) \exp\left(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt \quad (4.26)$$

für alle $j \in \{1, \dots, N_{hi}(\tau) + 1\}$. Mit (4.24) und den Ableitungen (4.25) liegen nun sämtliche Komponenten vor, die zur Bestimmung von (4.21) benötigt werden.

4.2 Asymptotische Verteilung unter Hypothese

In diesem Abschnitt soll nun gezeigt werden, dass unter der Nullhypothese „ $\eta_{+1} = \eta_{-1} = 0$ “ tatsächlich gilt

$$\frac{1}{n} \bar{U}_\eta^\top(\hat{\theta}) \hat{\Sigma}_{\eta\eta}^{-1}(\hat{\theta}) \bar{U}_\eta(\hat{\theta}) \stackrel{\mathcal{D}}{\sim} \chi_2^2.$$

Dazu ist nachzuweisen, dass der Score-Vektor von \bar{L} unter H_0 in einer Umgebung um die wahren Parameterwerte gleichmäßig gegen eine normalverteilte Zufallsvariable mit Erwartungswert 0 und der Fisher-Information als Kovarianzmatrix konvergiert. Bei einer Zählprozess-Likelihood kann man sich dazu üblicherweise die Martingaleeigenschaft des Score-Vektors zu Nutze machen. Dies ist hier jedoch nicht ohne Weiteres möglich, da \bar{L} lediglich auf einer Approximation einer Zählprozess-Likelihood basiert. Die Martingaleeigenschaft ist in diesem Fall nicht mehr offensichtlich. Es lässt sich allerdings zeigen, dass \bar{L} im Wesentlichen der Dichtefunktion der Zeitspannen zwischen den Ratingwechseln entspricht. Diese Erkenntnis lässt sich dann nutzen um Erwartungswertvektor und Kovarianzmatrix des Score-Vektors zu bestimmen. Für den Nachweis der H_0 -Verteilung der Teststatistik werden zusätzlich zu den Bedingungen **(A1)** bis **(A5)** aus Kapitel 2.3 die folgenden Regularitätsannahmen benötigt:

- (A6)** Die Zensierungszeiten der Ratinghistorien sind Realisationen der positiven u.i.v. Zufallsvariable C_1, \dots, C_n . Verteilungs- und Dichtefunktion von C_i seien für $i = 1, \dots, n$ mit G bzw. g bezeichnet und hängen nicht von den Parametern des Modells ab. Ferner seien die Zensierungszeiten unabhängig von den Sprungzeiten.
- (A7)** $\bar{U}_{+1i}(\theta_{+1}^{(0)})$ und $\bar{U}_{-1i}(\theta_{-1}^{(0)})$ sind für alle $i = 1, \dots, n$ unkorreliert.
- (A8)** Es existiert eine kompakte Umgebung Θ um $\theta^{(0)} := (\theta_{+1}^{(0)\top}, \theta_{-1}^{(0)\top})^\top$, so dass $\tilde{\mathcal{I}}(\theta)$ für alle $\theta \in \Theta$ positiv definit und damit invertierbar ist.
- (A9)** $\hat{\mathcal{I}}(\hat{\theta})$ aus Lemma 4.4 ist positiv definit und damit invertierbar.

Annahme **(A6)** stellt sicher, dass sich die Verteilung der Zensierungszeiten bei der Parameterinferenz vernachlässigen lässt. Ein offensichtlicher Grund für eine Verletzung dieser Annahme in der Praxis liegt nicht vor. Annahme **(A7)** ermöglicht Upgrades und Downgrades in den folgenden Beweisen getrennt voneinander zu betrachten und dient somit der Vereinfachung. Ein formaler Nachweis dieser Annahme gelingt nicht ohne weiteres. Intuitiv ist dies jedoch klar, da sich die Likelihood in einen Upgrade- und einen Downgradefaktor aufspalten lässt und die beiden Faktoren keine gemeinsamen Parameter aufweisen. Der Likelihoodfaktor für die Upgrades liefert somit keinen Beitrag zur Schätzung der Downgradeparameter und umgekehrt, so dass sich die entsprechenden Komponenten

des Score-Vektors unkorreliert sind. Annahme **(A8)** kann als erfüllt betrachtet werden, da die einzige Möglichkeit zur Verletzung dieser Annahme in der Singularität von $\tilde{\mathcal{I}}$ besteht. Dies scheint nach einem Blick auf die der Informationsmatrix zu Grunde liegenden Elemente in (4.25) jedoch nahezu unmöglich. Für einen formalen Nachweis dieser Aussage benötigt man allerdings die genaue Verteilung der $Y_{hi}(t)$ und der $\mathbf{X}_i(t)$. Der Schätzer für die Informationsmatrix sollte bei Gültigkeit von **(A8)** - zumindest für hinreichend große n - ebenfalls positiv definit sein. Dies war in Simulationsstudien für $n > 25$ auch immer der Fall, so dass Annahme **(A9)** als erfüllt betrachtet werden kann.

Sei im Folgenden

$$\bar{U}_i(\boldsymbol{\theta}) := \left. \frac{\partial}{\partial \boldsymbol{\theta}^*} \log \bar{L}_i(\boldsymbol{\theta}^*) \right|_{(\boldsymbol{\theta}_{+1}^\top, 0, \boldsymbol{\theta}_{-1}^\top, 0)}$$

bzw.

$$\bar{U}_{hi}(\boldsymbol{\theta}_h) := \left. \frac{\partial}{\partial \boldsymbol{\theta}^*} \log \bar{L}_{hi}(\boldsymbol{\theta}^*) \right|_{(\boldsymbol{\theta}_h^\top, 0)}.$$

Dabei bezeichnet \bar{L}_{hi} den Likelihoodfaktor von Schuldner i und Sprungtyp h , d.h.

$$\begin{aligned} \bar{L}_{hi}(\boldsymbol{\theta}_h^*) &= L_{hi}(\boldsymbol{\theta}_h) \left[1 + \eta_h z_{hi(N_{hi}(\tau)+1)}^2(\boldsymbol{\theta}_h) \right] \\ &\quad \times \prod_{j=1}^{N_{hi}(\tau)} \left[1 + \eta_h (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) \right], \end{aligned} \quad (4.27)$$

wobei $\boldsymbol{\theta}_h^* = (\boldsymbol{\theta}_h^\top, \eta_h)$ und

$$z_{hij}(\boldsymbol{\theta}_h) := \int_{T_{hi(j-1)}}^{T_{hij}} Y_{hi}(t) \lambda_{0h} \exp\left(\boldsymbol{\beta}_h^\top \mathbf{X}_i(t)\right) dt \quad (4.28)$$

für $j = 1, \dots, N_{hi}(\tau) + 1$.

Lemma 4.1 Die Annahme **(A6)** sei erfüllt. Für $i = 1, \dots, n$ und $h \in \{+1, -1\}$ gilt unter $H_0 : \eta_h = 0$

$$E[\bar{U}_{hi}(\boldsymbol{\theta}_h^{(0)})] = 0,$$

wobei $\boldsymbol{\theta}_h^{(0)} = (\lambda_{0h}^{(0)}, \boldsymbol{\beta}_h^{(0)\top})^\top$.

Beweis: Der Beweis unterscheidet sich für die Upgrades und Downgrades geringfügig (mehr dazu später) und wird lediglich für die Downgrades, also für $h = -1$, im Detail vorgestellt. Zunächst gilt es die gemeinsame (bedingte) Dichtefunktion der Zwischeneintrittszeiten der Downgrades, sowie der Zeit vom letzten Downgrade bis zur Zensierung unter Annahme der Intensitäten $\lambda_{hi}^*(t)$ aus (3.2) herzuleiten. Sei $i \in \{1, \dots, n\}$ und seien $D_{-1i1}, D_{-1i2} \dots$ die Zwischeneintrittszeiten der Downgrades. Die bedingte Verteilung von

D_{-1ij} , gegeben $(D_{-1i(j-1)}, \dots, D_{-1i1})^\top = (d_{-1i(j-1)}, \dots, d_{-1i1})^\top =: \mathbf{d}^{(j-1)}$, $\mathbf{X}_i(t) = \mathbf{x}_i(t)$ und $W_{-1i}(t) = w_{-1i}(t)$ ist für $j \in \mathbb{N}$ gegeben durch

$$\begin{aligned} & F_{D_{-1ij} | \mathbf{d}^{(j-1)}, \mathbf{x}_i(t), w_{-1i}(t)}(d_{-1ij}) \\ &= 1 - P(D_{-1ij} \geq d_{-1ij} | \mathbf{x}_i(t), w_{-1i}(t), D_{-1ik} = d_{-1ik}, k < j) \\ &= 1 - P\left(T_{-1ij} - \sum_{k=1}^{j-1} D_{-1ik} \geq d_{-1ij} \middle| \mathbf{x}_i(t), w_{-1i}(t), \sum_{k=1}^{j-1} D_{-1ik} = \sum_{k=1}^{j-1} d_{-1ik}\right). \end{aligned} \quad (4.29)$$

Dabei folgt die letzte Gleichheit in (4.29) mit $T_{-1ij} = \sum_{k=1}^j D_{-1ik}$ und der Annahme, dass es sich bei dem Rating eines Schuldners um einen Markov-Prozess handelt. Somit ist für die bedingte Verteilung von D_{-1ij} lediglich die Summe der vorherigen Zwischeneintrittszeiten relevant. Seien $\lambda_{-1i}^{(0)}(t | \mathbf{x}_i(t))$ und $\lambda_{-1i}^{(0)*}(t | \mathbf{x}_i(t), w_{-1i}(t))$ die Intensitäten (2.10) bzw. (3.2) an der Stelle der wahren Parameterwerte bei gegebenem Kovariablen- und Messfehlerprozess. Wegen $T_{-1ij} > T_{-1i(j-1)}$ lässt sich (4.29) weiter umformen zu

$$\begin{aligned} & F_{D_{-1ij} | \mathbf{d}^{(j-1)}, \mathbf{x}_i(t), w_{-1i}(t)}(d_{-1ij}) \\ &= 1 - P\left(T_{-1ij} - \sum_{k=1}^{j-1} d_{-1ik} \geq \sum_{k=1}^j d_{-1ik} - \sum_{k=1}^{j-1} d_{-1ik} \middle| \mathbf{x}_i(t), w_{-1i}(t), T_{-1i(j-1)} = \sum_{k=1}^{j-1} d_{-1ik}\right) \\ &= 1 - P\left(T_{-1ij} \geq \sum_{k=1}^j d_{-1ik} \middle| \mathbf{x}_i(t), w_{-1i}(t), T_{-1ij} > \sum_{k=1}^{j-1} d_{-1ik}\right) \\ &= 1 - \exp\left(-\int_{\sum_{k=1}^{j-1} d_{-1ik}}^{\sum_{k=1}^j d_{-1ik}} \lambda_{-1i}^{(0)*}(t | \mathbf{x}_i(t), w_{-1i}(t)) dt\right) \\ &= 1 - \exp\left(-v_{-1ij} \int_{\sum_{k=1}^{j-1} d_{-1ik}}^{\sum_{k=1}^j d_{-1ik}} \lambda_{-1i}^{(0)}(t | \mathbf{x}_i(t)) dt\right). \end{aligned} \quad (4.30)$$

Die letzte Zeile von (4.30) folgt dabei aus der Tatsache, dass der in $\lambda_{-1i}^{(0)*}(t | \mathbf{x}_i(t), w_{-1i}(t))$ auftauchende Pfad von $W_{-1i}(t)$ zwischen dem $(j-1)$ ten und j ten Downgrade stets einen konstanten Wert v_{-1ij} annimmt. Die zugehörige bedingte Dichte lautet

$$\begin{aligned} f_{D_{-1ij} | \mathbf{d}^{(j-1)}, \mathbf{x}_i(t), w_{-1i}(t)}(d_{-1ij}) &= v_{-1ij} \lambda_{-1i}^{(0)}\left(\sum_{k=1}^j d_{-1ik} \middle| \mathbf{x}_i(t)\right) \\ &\quad \times \exp\left(-v_{-1ij} \int_{\sum_{k=1}^{j-1} d_{-1ik}}^{\sum_{k=1}^j d_{-1ik}} \lambda_{-1i}^{(0)}(t | \mathbf{x}_i(t)) dt\right). \end{aligned} \quad (4.31)$$

Sei nun $\Delta_i \in \{0, 1\}$ ein Indikator, welcher angibt, ob das letzte an Schuldner i beobachtete Downgrade einen Kreditausfall darstellt. Im Fall $\Delta_i = 1$ werden lediglich die $N_{-1i}(\tau)$ Zwischeneintrittszeiten beobachtet. Da ein Kreditausfall jedoch ein seltenes Ereignis darstellt, nimmt Δ_i beim Großteil der Schuldner den Wert 0 an, was zur Folge hat, dass zusätzlich

noch die Zeit zwischen dem letzten Downgrade und der Zensierung beobachtet wird. Im Folgenden wird lediglich dieser Fall im Detail betrachtet.

Betrachte die Zeit zwischen dem n_{-1i} ten Downgrade und der Zensierung:

$$\tilde{C}_i := \left(C_i \left| C_i > \sum_{j=1}^{n_{-1i}} d_{-1ij} \right. \right) - \sum_{j=1}^{n_{-1i}} d_{-1ij}$$

Für die bedingte Dichte von C_i , gegeben $C_i > \sum_{j=1}^{n_{-1i}} d_{-1ij}$ an der Stelle c gilt wegen **(A7)**

$$\tilde{g}(c) := \frac{g(c)}{1 - G\left(\sum_{j=1}^{n_{-1i}} d_{-1ij}\right)} \mathbb{1}_{\left[\sum_{j=1}^{n_{-1i}} d_{-1ij}, \infty\right)}(c).$$

Somit ergibt sich die Dichte von \tilde{C}_i an der Stelle $c > 0$ zu $\tilde{g}(c + \sum_{j=1}^{n_{-1i}} d_{-1ij})$. Sei nun $N_{-1i}(\tau) = n_{-1i}$, $\mathbf{X}_i(t) = \mathbf{x}_i(t)$, $W_{-1i}(t) = w_{-1i}(t)$, $(D_{hin_{-1i}}, \dots, D_{hi1})^\top = \mathbf{d}^{(n_{-1i})}$ und $\Delta_i = 0$ gegeben. Dann gilt für die bedingte Dichte der Zufallsvariable $\tilde{D}_{-1i} = \min(\tilde{C}_i, D_{-1i(n_{-1i}+1)})$, welche die Zeitspanne zwischen dem n_{-1i} ten Downgrade und dem Minimum aus Zensierungszeitpunkt und nächstem Downgrade beschreibt:

$$\begin{aligned} & f_{\tilde{D}_{-1i} | \mathbf{d}^{(j)}, \mathbf{x}_i(t), w_{-1i}(t), n_{-1i}, \Delta_i=0}(\tilde{d}_{-1i}) \\ &= \frac{\partial}{\partial \tilde{d}_{-1i}} P\left(\tilde{D}_{-1i} \leq \tilde{d}_{-1i} \mid N_{-1i}(\tau) = n_{-1i}, \Delta_i = 0, \mathbf{d}^{(j-1)}, \mathbf{x}_i(t), w_{-1i}(t)\right) \quad (4.32) \\ &= \frac{\partial}{\partial \tilde{d}_{-1i}} P\left(\tilde{C}_i \leq \tilde{d}_{-1i}, \tilde{C}_i \leq D_{-1i(n_{-1i}+1)} \mid \mathbf{d}^{(j-1)}, \mathbf{x}_i(t), w_{-1i}(t)\right) \\ &= \frac{\partial}{\partial \tilde{d}_{-1i}} \int_0^{\tilde{d}_{-1i}} \int_v^\infty f_{D_{-1i(n_{-1i}+1)} | \mathbf{d}^{(n_{-1i})}, \mathbf{x}_i(t), w_{-1i}(t)}(u) \tilde{g}\left(v + \sum_{j=1}^{n_{-1i}} d_{-1ij}\right) dudv \\ &= \frac{\partial}{\partial \tilde{d}_{-1i}} \int_0^{\tilde{d}_{-1i}} \left(1 - F_{D_{-1i(n_{-1i}+1)} | \mathbf{d}^{(n_{-1i})}, \mathbf{x}_i(t), w_{-1i}(t)}(v)\right) \tilde{g}\left(v + \sum_{j=1}^{n_{-1i}} d_{-1ij}\right) dv \\ &= \left(1 - F_{\tilde{D}_{-1i} | \mathbf{d}^{(n_{hi})}, \mathbf{x}_i(t), w_{-1i}(t)}(\tilde{d}_{-1i})\right) \tilde{g}\left(\tilde{d}_{-1i} + \sum_{j=1}^{n_{-1i}} d_{-1ij}\right). \end{aligned}$$

Die gemeinsame bedingte Dichte von $\mathbf{D}_{-1i} := (D_{-1i1}, \dots, D_{-1in_{hi}})$ und \tilde{D}_{-1i} , gegeben $\mathbf{x}_i(t)$, n_{hi} und $\Delta_i = 0$, ergibt sich aus dem Produkt der einzelnen bedingten Dichten aus (4.31) und (4.32). Mit dieser gemeinsamen Dichte wäre es möglich, den Nachweis der Behauptung zu liefern. Als Vorbereitung auf den Beweis von Lemma 4.2, in dem es um die Kovarianz von $\bar{\mathbf{U}}_{hi}(\boldsymbol{\theta}_h^{(0)})$ geht, soll jedoch ein alternativer Weg eingeschlagen werden. Die bedingte Verteilungsfunktion (4.30) lässt sich nach dem selben Prinzip wie die Likelihood

in Kapitel 4.1 approximieren. Dabei erhält man

$$\begin{aligned}
 F_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t)}(d_{-1ij}) &\approx 1 - \exp \left[- \int_{\sum_{k=1}^{j-1} d_{-1ik}}^{\sum_{k=1}^j d_{-1ik}} \lambda_{-1i}^{(0)}(t|\mathbf{x}_i(t)) dt \right] \\
 &\times \left[1 + \eta_{-1} \left(\int_{\sum_{k=1}^{j-1} d_{-1ik}}^{\sum_{k=1}^j d_{-1ik}} \lambda_{-1i}^{(0)}(t|\mathbf{x}_i(t)) dt \right)^2 \right] \\
 &=: \bar{F}_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t)}(d_{-1ij}).
 \end{aligned} \tag{4.33}$$

Für $\eta_{-1} \in [0, 1]$ ist \bar{F} linksstetig mit 0 und 1 als links- bzw. rechtsseitigem Grenzwert und somit eine reguläre Verteilungsfunktion im Sinne der Definition von Kolmogoroff (vgl. Fisz (1989), Definition 2.2.1). Die zugehörige bedingte Dichte lautet

$$\begin{aligned}
 \bar{f}_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t)}(d_{-1ij}) &:= \lambda_{-1i}^{(0)} \left(\sum_{k=1}^j d_{hik} \middle| \mathbf{x}_i(t) \right) \exp \left[- \int_{\sum_{k=1}^{j-1} d_{hik}}^{\sum_{k=1}^j d_{hik}} \lambda_{-1i}^{(0)}(t|\mathbf{x}_i(t)) dt \right] \\
 &\times \left[1 + \eta_{-1} \left\{ \left(\int_{\sum_{k=1}^{j-1} d_{hik}}^{\sum_{k=1}^j d_{hik}} \lambda_{-1i}^{(0)}(t|\mathbf{x}_i(t)) dt \right)^2 \right. \right. \\
 &\left. \left. - 2 \int_{\sum_{k=1}^{j-1} d_{hik}}^{\sum_{k=1}^j d_{hik}} \lambda_{-1i}^{(0)}(t|\mathbf{x}_i(t)) dt \right\} \right]
 \end{aligned} \tag{4.34}$$

und stellt entsprechend eine Approximation der bedingten Dichte $f_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t),w_{-1i}(t)}$ der Zwischeneintrittszeiten dar. Die gemeinsame bedingte Dichte von $\mathbf{D}_{-1i} := (D_{-1i1}, \dots, D_{-1in_{hi}})$ und \tilde{D}_{-1i} , gegeben $\mathbf{x}_i(t)$, $w_{-1i}(t)$, n_{hi} und $\Delta_i = 0$ lässt sich nun mit Hilfe von (4.33) und (4.34) approximieren. Es gilt

$$\begin{aligned}
 &f_{\mathbf{D}_{-1i},\tilde{D}_{-1i}|\mathbf{x}_i(t),w_{-1i}(t),n_{hi},\Delta_i=0}(\mathbf{d}_{-1i},\tilde{d}_{-1i}) \\
 &= \left(\prod_{j=1}^{n_{hi}} f_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t),w_{-1i}(t)}(d_{-1ij}) \right) \\
 &\quad \times \left(1 - F_{\tilde{D}_{-1i}|\mathbf{d}^{(n_{hi})},\mathbf{x}_i(t),w_{-1i}(t)}(\tilde{d}_{-1i}) \right) \tilde{g} \left(\tilde{d}_{-1i} + \sum_{j=1}^{n_{-1i}} d_{-1ij} \right)
 \end{aligned} \tag{4.35}$$

$$\begin{aligned}
 &\approx \left(\prod_{j=1}^{n_{hi}} \bar{f}_{D_{-1ij}|\mathbf{d}^{(j-1)},\mathbf{x}_i(t)}(d_{-1ij}) \right) \\
 &\quad \times \left(1 - \bar{F}_{\tilde{D}_{-1i}|\mathbf{d}^{(n_{hi})},\mathbf{x}_i(t)}(\tilde{d}_{-1i}) \right) \tilde{g} \left(\tilde{d}_{-1i} + \sum_{j=1}^{n_{-1i}} d_{-1ij} \right)
 \end{aligned} \tag{4.36}$$

$$=: \bar{f}_{-1i}(\mathbf{d}_{-1i},\tilde{d}_{-1i};\eta_{-1}).$$

Unter der Nullhypothese $\eta_{-1} = 0$ nehmen die Messfehlergrößen v_{-1ij} für alle j den Wert 1 an. In diesem Fall entsprechen \bar{F} und \bar{f} aus (4.33) und (4.34) der tatsächlichen bedingten

Verteilungs- und Dichtefunktion aus (4.30) und (4.31). Unter H_0 sind (4.35) und (4.36) somit gleich.

Die Ereigniszeiten ergeben sich aus der Summe der Zwischeneintrittszeiten, d.h. es gilt $T_{-1ij} = \sum_{k=1}^j D_{-1ik}$ für alle $j \in [1, \dots, n_{-1i}]$. Außerdem ist $Y_{-1i}(t) = 1$ für alle $t \in [0, C_i]$. Mit (4.27) folgt somit, dass bei gegebenen Ereigniszeiten, Kovariablenverlauf und Risikoprozess gilt, dass

$$\bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^{(0)}, \eta_{-1}) = \frac{\bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1})}{\tilde{g}(\tilde{d}_{-1i} + \sum_{j=1}^{n_{-1i}} d_{-1ij})}. \quad (4.37)$$

Die Likelihood \bar{L}_{-1i} entspricht also bis auf den Faktor $\tilde{g}(\tilde{d}_{-1i} + \sum_{j=1}^{n_{-1i}} d_{-1ij})$ einer bedingten Dichte von Zwischeneintrittszeiten. Ferner gilt

$$\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} = \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \log [\bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1})], \quad (4.38)$$

d.h. Log-Likelihood und bedingte Dichte haben die gleiche Ableitung, da g nach **(A6)** nicht von $(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top$ abhängt. Aus der Tatsache, dass es sich bei (4.36) um eine (bedingte) Dichte handelt, folgt

$$\int_0^\infty \dots \int_0^\infty \bar{f}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} = 1. \quad (4.39)$$

Die Integrationsgrenzen hängen nicht von $(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top$ ab, d.h. Integration und Differentiation nach $(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top$ lassen sich vertauschen. Somit folgt mit Hilfe von (4.37) durch Ableiten beider Seiten von (4.39):

$$\begin{aligned} & \int_0^\infty \dots \int_0^\infty \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \\ &= \int_0^\infty \dots \int_0^\infty \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \log \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) \\ & \quad \times \bar{f}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \\ &= \int_0^\infty \dots \int_0^\infty \left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \right) \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \\ &= 0 \end{aligned} \quad (4.40)$$

Wie bereits erwähnt, entspricht \bar{f}_{-1i} für $\eta_{-1} = 0$ der tatsächlichen bedingten Dichte der Zwischeneintrittszeiten und der Zeit bis zur Zensurierung. Unter H_0 steht in der vorletzten Zeile von (4.40) also der bedingte Erwartungswert der Log-Likelihood, d.h. es gilt

$$\mathbb{E} \left[\left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \right) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)^\top} \Big| N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i = 0 \right] = 0. \quad (4.41)$$

Gleichung (4.41) gilt ebenfalls für $\Delta_i = 1$, also für den (seltenen) Fall, dass die Ratinghistorie eines Schuldners mit einem Downgrade in die Insolvenzklasse endet. Der Nachweis gelingt analog. Der einzige Unterschied besteht darin, dass die Zeit zwischen dem letzten Downgrade und der Zensierung in diesem Fall nicht berücksichtigt wird und somit der letzte Faktor von (4.36) entfällt. Demnach gilt

$$E [\bar{U}_{-1i}(\boldsymbol{\theta}_{-1}^{(0)}) | N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i] = 0. \quad (4.42)$$

Nach dem Satz vom iterierten Erwartungswert verschwindet der Erwartungswert auch im unbedingten Fall. Damit folgt die Behauptung für $h = -1$.

Für den Fall $h = +1$, also für die einstufigen Upgrades, lässt sich der Beweis analog durchführen. Hier ist allerdings der Spezialfall zu berücksichtigen, dass der Risikoprozess $Y_{+1i}(t)$ nicht nur zum Zeitpunkt der Zensierung, sondern auch zwischenzeitlich den Wert 0 annehmen kann. Dies ist der Fall, wenn sich Schuldner i in der höchstmöglichen Ratingklasse befindet, was zur Folge hat, dass er nicht unter dem „Risiko“ einer weiteren Ratingverbesserung steht. Ein weiterer Unterschied zu den Downgrades ist die Tatsache, dass die Beobachtungszeit eines Schuldners nicht mit einem Upgrade enden kann, sondern stets eine positive Zeitspanne zwischen dem letzten Upgrade und dem Zensierungszeitpunkt beobachtet wird. \square

Lemma 4.2 *Die Annahme (A6) sei erfüllt. Für $i = 1, \dots, n$ und $h \in \{+1, -1\}$ gilt unter $H_0 : \eta_h = 0$*

$$\text{Cov} [\bar{U}_{hi}(\boldsymbol{\theta}_h^{(0)})] = -E \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}_h^* (\partial \boldsymbol{\theta}_h^*)^\top} \log \bar{L}_{hi}(\boldsymbol{\theta}_h^*) \right)_{(\boldsymbol{\theta}_h^{(0)\top}, 0)^\top} \right] =: \mathcal{I}_h(\boldsymbol{\theta}_h^{(0)}).$$

Beweis: Auch in diesem Beweis wird lediglich der Fall $h = -1$ und $\Delta_i = 0$ betrachtet. Erneutes Ableiten von (4.40) nach $(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top$ mit Hilfe der Produktregel liefert

$$\begin{aligned} 0 &= \int_0^\infty \dots \int_0^\infty \frac{\partial^2}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top \partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})} \log \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) \\ &\quad \times \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \\ &\quad + \int_0^\infty \dots \int_0^\infty \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \log \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) \\ &\quad \times \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})} \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1}. \end{aligned} \quad (4.43)$$

Mit

$$\begin{aligned} \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})} \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) &= \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) \\ &\times \frac{\partial}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})} \log \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}), \end{aligned}$$

Gleichung (4.38) und

$$\begin{aligned} &\left. \frac{\partial^2}{\partial \boldsymbol{\theta}_{-1}^* \partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \right|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \\ &= \frac{\partial^2}{\partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top \partial (\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})} \log \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) \end{aligned}$$

folgt somit aus (4.43):

$$\begin{aligned} &\int_0^\infty \dots \int_0^\infty \frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \right)^\top \\ &\quad \times \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \\ &= - \int_0^\infty \dots \int_0^\infty \frac{\partial^2}{\partial \boldsymbol{\theta}_{-1}^* \partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \Big|_{(\boldsymbol{\theta}_{-1}^{(0)\top}, \eta_{-1})^\top} \\ &\quad \times \bar{f}_{-1i}(\mathbf{d}_{-1i}, \tilde{d}_{-1i}; \eta_{-1}) d\tilde{d}_{-1i} \dots dd_{-1i1} \end{aligned} \quad (4.44)$$

Unter H_0 entspricht \bar{f}_{-1i} der wahren bedingten Dichte von Zwischeneintrittszeiten und Zeit vom letztem Downgrade bis zur Zensierung. Gleichung (4.44) besagt also nichts anderes als

$$\begin{aligned} &\mathbb{E} \left[\left(\left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \right)^\top \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \Big| N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i = 0 \right] \\ &= - \mathbb{E} \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}_{-1}^* \partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i}(\boldsymbol{\theta}_{-1}^*) \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \Big| N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i = 0 \right]. \end{aligned} \quad (4.45)$$

Das gleiche Resultat lässt sich für den Fall $\Delta_i = 1$ erzielen. Insgesamt gilt mit Lemma 4.1, dem Satz vom iterierten Erwartungswert und Gleichung (4.45)

$$\begin{aligned}
 & \text{Cov} [\bar{U}_{-1i}(\boldsymbol{\theta}_{-1}^{(0)})] \\
 &= \text{E} \left[\left(\left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i} \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i} \right)^\top \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \right] \\
 &= \text{E} \left[\text{E} \left[\left(\left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^*} \log \bar{L}_{-1i} \right) \left(\frac{\partial}{\partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i} \right)^\top \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \middle| N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i \right] \right] \\
 &= - \text{E} \left[\text{E} \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}_{-1}^* \partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i} \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \middle| N_{-1i}(\tau), \mathbf{X}_i(t), \Delta_i \right] \right] \\
 &= - \text{E} \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}_{-1}^* \partial \boldsymbol{\theta}_{-1}^{*\top}} \log \bar{L}_{-1i} \right)_{(\boldsymbol{\theta}_{-1}^{(0)\top}, 0)} \right].
 \end{aligned}$$

Mit den im Beweis von Lemma 4.1 angesprochenen Modifikationen folgt die Behauptung für $h = +1$ analog. \square

Nach Lemma 4.1 und 4.2 sind nun Erwartungswert und Kovarianzmatrix von \bar{U}_{hi} an der Stelle $\boldsymbol{\theta}_h^{(0)}$ bekannt. Man beachte, dass diese beiden Größen nicht von i abhängen. Da die Prozesse $N_{hi}(t)$, $Y_{hi}(t)$ und $\mathbf{X}_i(t)$ für $i = 1, \dots, n$ als unabhängig und identisch verteilt angenommen wurden, gilt dies auch für die Zufallsvariablen \bar{U}_{hi} . Diese Resultate lassen sich nun ausnutzen, um die asymptotische Verteilung von $\bar{U} = \sum_{i=1}^n \bar{U}_i = \sum_{i=1}^n (\bar{U}_{+1i}, \bar{U}_{-1i})^\top$ zu bestimmen.

Lemma 4.3 *Die Annahmen (A6) und (A7) seien erfüllt. Sei Θ eine kompakte Umgebung um $\boldsymbol{\theta}^{(0)} := (\boldsymbol{\theta}_{+1}^{(0)\top}, \boldsymbol{\theta}_{-1}^{(0)\top})^\top$ und seien $\bar{\mathbf{u}}(\boldsymbol{\theta}) := E[\bar{U}_1(\boldsymbol{\theta})]$ und $\tilde{\mathcal{I}}(\boldsymbol{\theta}) := \text{Cov}[\bar{U}_1(\boldsymbol{\theta})]$ für $\boldsymbol{\theta} \in \Theta$. Dann gilt*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\bar{U}_i(\boldsymbol{\theta}) - \bar{\mathbf{u}}(\boldsymbol{\theta})) \underset{\Theta}{\overset{\mathcal{D}}{\rightsquigarrow}} N(\mathbf{0}, \tilde{\mathcal{I}}(\boldsymbol{\theta})).$$

Dabei bezeichnet „ $\underset{\Theta}{\overset{\mathcal{D}}{\rightsquigarrow}}$ “ die gleichmäßige Konvergenz in Verteilung für alle $\boldsymbol{\theta} \in \Theta$ (vgl. Borovkov (1998), S. 220). Daraus folgt insbesondere, dass

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \bar{U}_i(\boldsymbol{\theta}^{(0)}) \overset{\mathcal{D}}{\rightsquigarrow} N(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta}^{(0)})), \tag{4.46}$$

wobei

$$\mathcal{I}(\boldsymbol{\theta}^{(0)}) = -E \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}^* \partial \boldsymbol{\theta}^{*\top}} \log \bar{L}_i \right)_{(\boldsymbol{\theta}^{(0)}, 0)} \right]$$

mit $i \in \{1, \dots, n\}$ beliebig.

Beweis: Seien $i \in \{1, \dots, n\}$ und $\boldsymbol{\theta} \in \Theta$ beliebig. Betrachte $f_{\boldsymbol{\theta}}(N_{hi}(t), Y_{hi}(t), \mathbf{X}_i(t), h \in \{+1, -1\}, t \in [0, \tau]) = \bar{\mathbf{U}}_i(\boldsymbol{\theta})$, also den Score-Vektor von Schuldner i aufgefasst als Funktion der zugrunde liegenden Zufallsvariablen. Die Komponenten von $\bar{\mathbf{U}}_i(\boldsymbol{\theta})$ sind gegeben durch

$$\begin{aligned} \frac{\partial}{\partial \lambda_{0h}} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\boldsymbol{\theta}_{+1,0}^\top, \boldsymbol{\theta}_{-1,0}^\top)^\top} &= \frac{N_{hi}(\tau)}{\tilde{\lambda}_{0h}} - \int_0^\tau Y_{hi}(t) \exp\left(\tilde{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt, \\ \frac{\partial}{\partial \beta_l} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\boldsymbol{\theta}_{+1,0}^\top, \boldsymbol{\theta}_{-1,0}^\top)^\top} &= \sum_{j=1}^{N_{hi}(\tau)} X_{il}(T_{hij}) - \int_0^\tau Y_{hi}(t) \tilde{\lambda}_{0h} X_{il}(t) \exp\left(\tilde{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt, \end{aligned}$$

$l = 1, \dots, d$, und

$$\frac{\partial}{\partial \eta_h} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\boldsymbol{\theta}_{+1,0}^\top, \boldsymbol{\theta}_{-1,0}^\top)^\top} = \sum_{i=1}^n \left(\sum_{j=1}^{N_{hi}(\tau)} (z_{hij}^2(\boldsymbol{\theta}_h) - 2z_{hij}(\boldsymbol{\theta}_h)) + z_{hi(N_{hi}(\tau)+1)}^2(\boldsymbol{\theta}_h) \right)$$

mit

$$z_{hij}(\boldsymbol{\theta}_h) = \int_{T_{hi(j-1)}}^{T_{hij}} Y_{hi}(t) \tilde{\lambda}_{0h} \exp\left(\tilde{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt.$$

$f_{\boldsymbol{\theta}} = \bar{\mathbf{U}}_i(\boldsymbol{\theta})$ ist also offensichtlich beliebig oft differenzierbar bezüglich $\boldsymbol{\theta}$. Außerdem ist $f_{\boldsymbol{\theta}}$ aufgrund der Beschränktheit der zugrunde liegenden Zufallsvariablen beschränkt. Damit handelt es sich bei $\{f_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\}$ nach van der Vaart (1998), Beispiel 19.9, um eine *Donsker*-Menge. Aufgrund der Tatsache, dass die Zufallsvariablen $\bar{\mathbf{U}}_i(\boldsymbol{\theta})$ für $i = 1, \dots, n$ stochastisch unabhängig sind, folgt nach van der Vaart (1998), S. 269, dass

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\bar{\mathbf{U}}_i(\boldsymbol{\theta}) - \bar{\mathbf{u}}(\boldsymbol{\theta})) \stackrel{\mathcal{D}}{\underset{\Theta}{\rightsquigarrow}} N(\mathbf{0}, \bar{\mathcal{I}}(\boldsymbol{\theta})).$$

Aus der gleichmäßigen Konvergenz in Verteilung folgt die punktweise Konvergenz, d.h. es gilt insbesondere

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\bar{\mathbf{U}}_i(\boldsymbol{\theta}^{(0)}) - \bar{\mathbf{u}}_i(\boldsymbol{\theta}^{(0)})) \stackrel{\mathcal{D}}{\rightsquigarrow} N(\mathbf{0}, \bar{\mathcal{I}}(\boldsymbol{\theta}^{(0)})).$$

Nach Lemma 4.1 ist $\bar{\mathbf{u}}_i(\boldsymbol{\theta}^{(0)}) = \mathbf{0}$. Aus Lemma 4.2 folgt zusammen mit **(A7)**, dass

$$\bar{\mathcal{I}}(\boldsymbol{\theta}^{(0)}) = \begin{pmatrix} \mathcal{I}_{+1}(\boldsymbol{\theta}^{(0)}) & \mathbf{0} \\ \mathbf{0} & \mathcal{I}_{-1}(\boldsymbol{\theta}^{(0)}) \end{pmatrix}.$$

Aufgrund der Tatsache, dass die bezüglich Upgrade- und Downgradparametern gemischten zweiten Ableitungen von $\log \bar{L}_i$ verschwinden, gilt $\bar{\mathcal{I}}(\boldsymbol{\theta}^{(0)}) = \mathcal{I}(\boldsymbol{\theta}^{(0)})$. \square

Lemma 4.4 Die Annahmen **(A1)** bis **(A6)** seien erfüllt. Dann gilt unter H_0 :

$$\hat{\mathcal{I}}(\hat{\boldsymbol{\theta}}) = -\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^* (\partial \boldsymbol{\theta}^*)^\top} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\hat{\boldsymbol{\theta}}_{+1,0}^\top, \hat{\boldsymbol{\theta}}_{-1,0}^\top)} \xrightarrow{\mathcal{P}} \mathcal{I}(\boldsymbol{\theta}^{(0)}),$$

d.h. $\hat{\mathcal{I}}(\hat{\boldsymbol{\theta}})$ ist ein konsistenter Schätzer für $\mathcal{I}(\boldsymbol{\theta}^{(0)})$.

Beweis: Die Behauptung lässt sich analog zum Beweis von Satz 2.4 mit Hilfe von Davidson (1994) zeigen. Sei Θ eine kompakte Umgebung um $\boldsymbol{\theta}^{(0)}$. Nach dem Gesetz der großen Zahl gilt für alle $\boldsymbol{\theta} = (\boldsymbol{\theta}_{+1}^\top, \boldsymbol{\theta}_{-1}^\top)^\top \in \Theta$

$$\hat{\mathcal{I}}(\boldsymbol{\theta}) = -\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \boldsymbol{\theta}^* (\partial \boldsymbol{\theta}^*)^\top} \log \bar{L}_i(\boldsymbol{\theta}^*) \Big|_{(\boldsymbol{\theta}_{+1,0}^\top, \boldsymbol{\theta}_{-1,0}^\top)} \xrightarrow{\mathcal{P}} \mathcal{I}(\boldsymbol{\theta}),$$

wobei

$$\mathcal{I}(\boldsymbol{\theta}) := \mathbb{E} \left[\left(\frac{\partial^2}{\partial \boldsymbol{\theta}^* (\partial \boldsymbol{\theta}^*)^\top} \log \bar{L}_i(\boldsymbol{\theta}^*) \right)_{(\boldsymbol{\theta}_{+1,0}^\top, \boldsymbol{\theta}_{-1,0}^\top)} \right].$$

Aus (4.25) geht hervor, dass die einzelnen Einträge von $\hat{\mathcal{I}}(\boldsymbol{\theta})$ differenzierbar bezüglich $\boldsymbol{\theta}$ sind. Zusammen mit der Beschränktheit der in $\hat{\mathcal{I}}(\boldsymbol{\theta})$ auftauchenden Zufallsvariablen, sowie von Θ folgt mit der selben Argumentation wie im Beweis von Satz 2.4 mit Hilfe des Mittelwertsatzes und der Cauchy-Schwarz Ungleichung, dass die Einträge von $\hat{\mathcal{I}}(\boldsymbol{\theta})$ für alle $\boldsymbol{\theta} \in \Theta$ gleichmäßig gegen die Einträge von $\mathcal{I}(\boldsymbol{\theta})$ konvergieren. Da unter H_0 nach Satz 2.1 und Satz 2.4 außerdem gilt, dass $\hat{\boldsymbol{\theta}} \xrightarrow{\mathcal{P}} \boldsymbol{\theta}$, folgt mit Davidson (1994), Satz 21.6 die Behauptung. \square

Ersetzt man $\boldsymbol{\theta}^{(0)}$ und $\mathcal{I}(\boldsymbol{\theta}^{(0)})$ in (4.46) durch die konsistenten Schätzer $\hat{\boldsymbol{\theta}}$ und $\hat{\mathcal{I}}(\hat{\boldsymbol{\theta}})$, ergibt sich ein Problem bezüglich der Ableitungen der log-Likelihood nach $\lambda_{0(+1)}$ und $\lambda_{0(-1)}$. Im Hinblick auf die Notation aus Kapitel 4.1 seien diese beiden Komponenten von $\bar{\mathbf{U}}$ mit $\bar{\mathbf{U}}_{\lambda_0}$ bezeichnet. Bei $\hat{\lambda}_0$ handelt es sich um den ML-Schätzer für λ_0 im Modell aus Kapitel 2.2, d.h.

$$\frac{\partial}{\partial \lambda_0} \log L(\boldsymbol{\theta}) \Big|_{\hat{\lambda}_0, \hat{\boldsymbol{\beta}}} = 0.$$

Aus Gleichung (4.22) folgt

$$\frac{\partial}{\partial \boldsymbol{\theta}} \log L(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log \bar{L}(\boldsymbol{\theta}) \Big|_{\boldsymbol{\eta}=0}.$$

Somit ist $\bar{\mathbf{U}}_{\lambda_0}(\hat{\boldsymbol{\theta}}) = 0$, d.h. (4.46) bleibt nicht erhalten wenn man $\boldsymbol{\theta}^{(0)}$ durch $\hat{\boldsymbol{\theta}}$ ersetzt. Da sich $\hat{\boldsymbol{\beta}}$ aus einer partiellen Likelihood ergibt und $\boldsymbol{\eta}$ durch die Nullhypothese festgelegt

wird, ergeben sich diese Probleme in den zugehörigen Komponenten von \bar{U} - im Folgenden mit $\bar{U}_{-\lambda_0}$ bezeichnet - nicht. Allerdings sind \bar{U}_{λ_0} und $\bar{U}_{-\lambda_0}$ nicht unabhängig, so dass die unschönen Eigenschaften von \bar{U}_{λ_0} Auswirkungen auf $\bar{U}_{-\lambda_0}$ haben. Daher basiert die Teststatistik nicht auf $\bar{U}_{-\lambda_0}$ selbst, sondern auf einer um den Einfluss von \bar{U}_{λ_0} bereinigten Modifikation.

Lemma 4.5 *Die Annahmen (A6) bis (A8) seien erfüllt. Sei $\theta \in \Theta$ beliebig. Dann sind $\bar{U}_{\lambda_0}(\theta)$ und*

$$\bar{U}_{-\lambda_0}^*(\theta) = \bar{U}_{-\lambda_0}(\theta) - \tilde{\mathcal{I}}_{-\lambda_0\lambda_0}(\theta)\tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\bar{U}_{\lambda_0}(\theta)$$

unkorreliert.

Beweis:

$$\begin{aligned} \text{Cov} [\bar{U}_{\lambda_0}(\theta), \bar{U}_{-\lambda_0}^*(\theta)] &= \text{Cov} [\bar{U}_{\lambda_0}(\theta), \bar{U}_{-\lambda_0}(\theta) - \tilde{\mathcal{I}}_{-\lambda_0\lambda_0}(\theta)\tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\bar{U}_{\lambda_0}(\theta)] \\ &= \text{Cov} [\bar{U}_{\lambda_0}(\theta), \bar{U}_{-\lambda_0}(\theta)] \\ &\quad - \text{Cov} [\bar{U}_{\lambda_0}(\theta), \tilde{\mathcal{I}}_{-\lambda_0\lambda_0}(\theta)\tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\bar{U}_{\lambda_0}(\theta)] \\ &= \tilde{\mathcal{I}}_{\lambda_0-\lambda_0}(\theta) - \text{Cov} [\bar{U}_{\lambda_0}(\theta), \bar{U}_{\lambda_0}(\theta)] \tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\tilde{\mathcal{I}}_{-\lambda_0\lambda_0}^\top(\theta) \\ &= \tilde{\mathcal{I}}_{\lambda_0-\lambda_0}(\theta) - \tilde{\mathcal{I}}_{\lambda_0\lambda_0}(\theta)\tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\tilde{\mathcal{I}}_{\lambda_0-\lambda_0}(\theta) \\ &= 0 \end{aligned} \tag{4.47}$$

□

Satz 4.1 *Die Annahmen (A1) bis (A9) seien erfüllt. Sei*

$$\hat{\Sigma}(\hat{\theta}) := \hat{\mathcal{I}}_{-\lambda_0-\lambda_0}(\hat{\theta}) - \hat{\mathcal{I}}_{-\lambda_0\lambda_0}(\hat{\theta})\hat{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\hat{\theta})\hat{\mathcal{I}}_{\lambda_0-\lambda_0}(\hat{\theta})$$

und seien $\hat{\Sigma}_{\eta\eta}(\hat{\theta}) \in \mathbb{R}^{2 \times 2}$ diejenigen Zeilen und Spalten von $\hat{\Sigma}(\hat{\theta})$, die in $\hat{\mathcal{I}}_{-\lambda_0-\lambda_0}(\hat{\theta})$ zu den zweiten Ableitungen bezüglich η gehören. Dann gilt unter der Nullhypothese „ $\eta = \mathbf{0}$ “:

$$\frac{1}{n}\bar{U}_\eta^\top(\hat{\theta})\hat{\Sigma}_{\eta\eta}^{-1}(\hat{\theta})\bar{U}_\eta(\hat{\theta}) \stackrel{\mathcal{D}}{\sim} \chi_2^2.$$

Beweis: Sei $\theta \in \Theta$ und $\bar{u}_{-\lambda_0}^*(\theta) := \mathbb{E}[\bar{U}_{-\lambda_0}^*(\theta)]$. Da es sich bei $\bar{U}_{-\lambda_0}^*$ um eine lineare Transformation von $\bar{U}_{-\lambda_0}$ handelt, gilt wegen Lemma 4.3

$$\frac{1}{\sqrt{n}}(\bar{U}_{-\lambda_0}^*(\theta) - n\bar{u}_{-\lambda_0}^*(\theta)) \stackrel{\mathcal{D}}{\underset{\Theta}{\sim}} N(\mathbf{0}, \Sigma(\theta))$$

mit

$$\Sigma(\theta) = \tilde{\mathcal{I}}_{-\lambda_0-\lambda_0}(\theta) - \tilde{\mathcal{I}}_{-\lambda_0\lambda_0}(\theta)\tilde{\mathcal{I}}_{\lambda_0\lambda_0}^{-1}(\theta)\tilde{\mathcal{I}}_{\lambda_0-\lambda_0}(\theta).$$

Nach Rao (2005) sind zur Überprüfung von $H_0 : \boldsymbol{\eta} = \mathbf{0}$ nur die Komponenten der Modifikation von $\bar{\mathbf{U}}_{-\lambda_0}$ relevant, die zu den Parametern $\boldsymbol{\eta}$ gehören. Diese seien im Folgenden mit $\bar{\mathbf{U}}_{\boldsymbol{\eta}}^*$ bezeichnet und der zugehörige Erwartungswertvektor entsprechend mit $\bar{\mathbf{u}}_{\boldsymbol{\eta}}^*$. Es gilt

$$\frac{1}{\sqrt{n}} (\bar{\mathbf{U}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta}) - n\bar{\mathbf{u}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta})) \stackrel{\mathcal{D}}{\underset{\Theta}{\approx}} N(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}(\boldsymbol{\theta})),$$

wobei $\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}(\boldsymbol{\theta}) \in \mathbb{R}^{2 \times 2}$ diejenigen Zeilen und Spalten von $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ bezeichnet, die in $\bar{\boldsymbol{\mathcal{I}}}_{-\lambda_0 - \lambda_0}(\boldsymbol{\theta})$ zu den zweiten Ableitungen bezüglich $\boldsymbol{\eta}$ gehören. Mit (A8) folgt nun

$$\frac{1}{\sqrt{n}} \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\boldsymbol{\theta}) (\bar{\mathbf{U}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta}) - n\bar{\mathbf{u}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta})) \stackrel{\mathcal{D}}{\underset{\Theta}{\approx}} N(\mathbf{0}, \mathbb{I}_2). \quad (4.48)$$

Mit Lemma 4.1 und Gleichung (4.47) ergibt sich unter der Nullhypothese

$$\bar{\mathbf{u}}_{-\lambda_0}^*(\boldsymbol{\theta}^{(0)}) = \mathbb{E} [\bar{\mathbf{U}}_{-\lambda_0}(\boldsymbol{\theta}^{(0)})] - \bar{\boldsymbol{\mathcal{I}}}_{-\lambda_0 \lambda_0}(\boldsymbol{\theta}^{(0)}) \bar{\boldsymbol{\mathcal{I}}}_{\lambda_0 \lambda_0}^{-1}(\boldsymbol{\theta}^{(0)}) \mathbb{E} [\bar{\mathbf{U}}_{\lambda_0}(\boldsymbol{\theta}^{(0)})] = 0$$

und somit folgt aus (4.48), dass

$$\frac{1}{\sqrt{n}} \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\boldsymbol{\theta}^{(0)}) \bar{\mathbf{U}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta}^{(0)}) \stackrel{\mathcal{D}}{\approx} N(\mathbf{0}, \mathbb{I}_2). \quad (4.49)$$

Aufgrund von (4.48) und Lemma 4.5 behält (4.49) Gültigkeit, wenn man $\boldsymbol{\theta}^{(0)}$ in $\bar{\mathbf{U}}_{\boldsymbol{\eta}}^*(\boldsymbol{\theta}^{(0)})$ durch den nach Satz 2.1 und Satz 2.4 unter H_0 konsistenten Schätzer $\hat{\boldsymbol{\theta}}$ ersetzt. Wegen $\bar{\mathbf{U}}_{\lambda_0}(\hat{\boldsymbol{\theta}}) = 0$ folgt allerdings in diesem Fall unmittelbar aus Gleichung (4.47), dass $\bar{\mathbf{U}}_{-\lambda_0}^*(\hat{\boldsymbol{\theta}}) = \bar{\mathbf{U}}_{-\lambda_0}(\hat{\boldsymbol{\theta}})$ und damit auch, dass $\bar{\mathbf{U}}_{\boldsymbol{\eta}}^*(\hat{\boldsymbol{\theta}}) = \bar{\mathbf{U}}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\theta}})$. Somit gilt

$$\frac{1}{\sqrt{n}} \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\boldsymbol{\theta}^{(0)}) \bar{\mathbf{U}}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\theta}}) \stackrel{\mathcal{D}}{\approx} N(\mathbf{0}, \mathbb{I}_2). \quad (4.50)$$

Aus Lemma 4.4 folgt zusammen mit (A9) und dem Continuous-Mapping-Theorem, dass $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\hat{\boldsymbol{\theta}}) \xrightarrow{\mathcal{P}} \boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\boldsymbol{\theta}^{(0)})$. Mit dem Satz von Slutsky ergibt sich schließlich, dass

$$\frac{1}{\sqrt{n}} \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\eta}\boldsymbol{\eta}}^{-1/2}(\hat{\boldsymbol{\theta}}) \bar{\mathbf{U}}_{\boldsymbol{\eta}}(\hat{\boldsymbol{\theta}}) \stackrel{\mathcal{D}}{\approx} N(\mathbf{0}, \mathbb{I}_2) \quad (4.51)$$

und damit folgt die Behauptung. \square

4.3 Güte und Niveau in endlichen Stichproben

Die Eigenschaften des Tests sollen nun anhand einer Simulationsstudie überprüft werden. Dazu werden dieselben Simulationsparameter wie in Kapitel 3.2 verwendet. Von Interesse sind die Eigenschaften der Teststatistik-Verteilung unter „ $H_0: \eta_h = 0 \forall h \in \{+1, -1\}$ “ und die Güte des Tests bei einem Signifikanzniveau von 5% in verschiedenen Stichprobengrößen. Die Studie wird mit einer maximalen Stichprobe von $n = 1000$ Schuldnern

durchgeführt. Sämtliche Simulationsergebnisse beruhen auf 5000 Iterationen. Abbildung 4.1 zeigt die geschätzte Verteilung bei Gültigkeit der Nullhypothese. Die Verteilung der Teststatistik scheint recht schnell gegen die χ_2^2 -Verteilung zu konvergieren. Bereits bei einer Stichprobengröße von $n = 250$ ergibt sich eine zufriedenstellende Annäherung. Tabelle 4.1 zeigt die Güte des Tests gegenüber Messfehlern bei einem Signifikanzniveau von 5%. Die Wahrscheinlichkeit für einen Fehler 1. Art liegt bei den betrachteten Stichprobengrößen noch über den geforderten 5% und nähert sich diesem Wert von oben an. Der Test entscheidet also antikonservativ. Bei $n = 1000$ wird das gewünschte Niveau jedoch in etwa erreicht. Insgesamt erweist sich der Test als recht trennscharf, da er selbst Messfehler mit einer kleinen Varianz bei einer überschaubaren Stichprobengröße von $n = 250$ zuverlässig erkennt.

Neben Messfehlern in den Ratingwechselzeitpunkten sind auch andere Ursachen für un beobachtbare Heterogenität in den Daten denkbar. Die Annahme, dass der Test lediglich auf Messfehler im Sinne des Modells (3.2) reagiert ist unrealistisch. Daher soll nun die Sensibilität des Tests gegenüber einer weiteren Form von latenter Datenheterogenität untersucht werden: vernachlässigte Kovariablen. Seien $\tilde{\mathbf{V}}_h = (\tilde{V}_{hi}, i \in \{1, \dots, n\})$, $h \in \{-1, +1\}$, zwei unabhängige Familien von unabhängig identisch verteilten positiven Zufallsvariablen

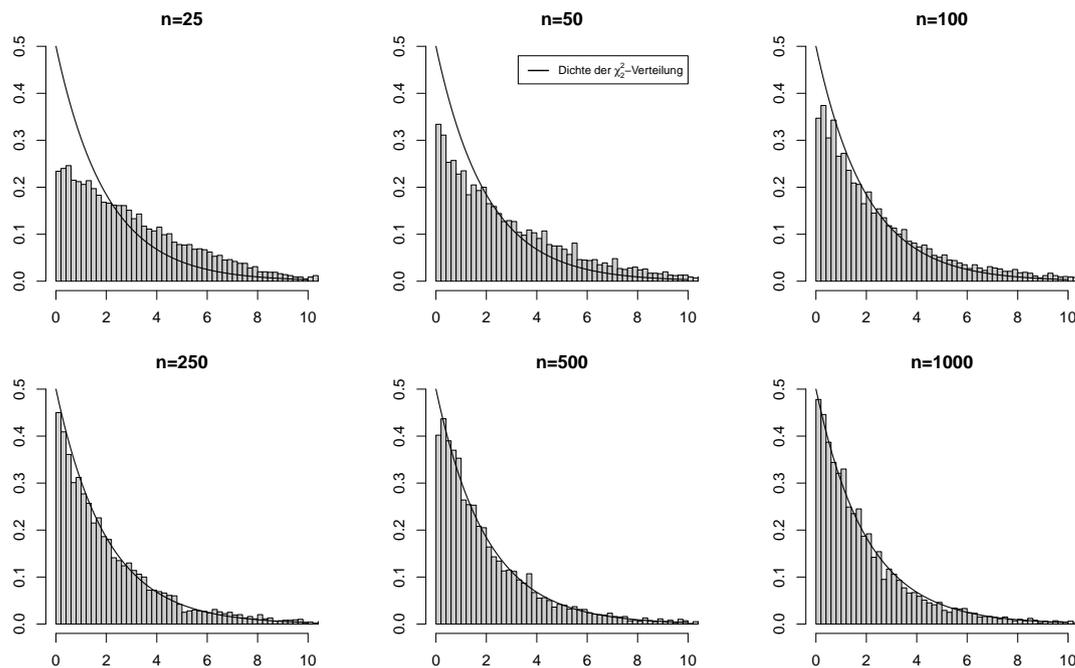


Abbildung 4.1: Verteilung der Teststatistik unter der Nullhypothese „ $\eta = 0$ “ für verschiedene Stichprobengrößen n

Tabelle 4.1: Güte des Tests bei Messfehlern ($\alpha = 5\%$)

Varianz	Stichprobengröße					
	25	50	100	250	500	1000
0.0	0.128	0.123	0.100	0.079	0.065	0.055
0.5	0.096	0.324	0.715	0.997	1.000	1.000
1.0	0.211	0.669	0.953	1.000	1.000	1.000
1.5	0.292	0.792	0.978	1.000	1.000	1.000
2.0	0.339	0.828	0.989	1.000	1.000	1.000
2.5	0.359	0.849	0.991	1.000	1.000	1.000

mit Erwartungswert 1. Ein Intensitätsmodell für Schuldner i unter Berücksichtigung von latenten zeitkonstanten Kovariablen ist gegeben durch

$$\lambda_{hi}^{**}(t, V_{hi}) = V_{hi} \lambda_{hi}(t), \quad h \in \{+1, -1\}. \quad (4.52)$$

Bezeichnet man die vernachlässigten Kovariablen mit $\tilde{\mathbf{X}}_i$ und die zugehörigen Einflussparameter mit $\tilde{\boldsymbol{\beta}}_h$, so gilt $\tilde{V}_{hi} = \exp(\tilde{\boldsymbol{\beta}}_h^\top \tilde{\mathbf{X}}_i)$. Man beachte, dass den Familien $\tilde{\mathbf{V}}_{+1}$ und $\tilde{\mathbf{V}}_{-1}$ nicht die gleiche Verteilung unterstellt wird, da sich $\tilde{\boldsymbol{\beta}}_{+1}$ und $\tilde{\boldsymbol{\beta}}_{-1}$ im Allgemeinen unterscheiden. Beiden Familien wird jedoch als Identifizierbarkeitsbedingung ein Erwartungswert von 1 unterstellt. Dies ist ohne Beschränkung der Allgemeinheit möglich, da Abweichungen von diesem Erwartungswert durch die Basisintensität λ_{0h} , welche als Faktor in $\lambda_h(t)$ auftaucht, absorbiert werden. Auch bei vernachlässigten Kovariablen stellt sich zunächst die Frage nach den Auswirkungen auf die Parameterschätzer, welche - exemplarisch für die Downgradeschätzer - durch Abbildung 4.2 beantwortet wird. Die zugrunde liegenden Simulationsparametern sind die gleichen wie in Kapitel 3.2. Die Auswirkungen auf die Schätzer für die Kovariableneinflüsse sind ähnlich wie bei der Vernachlässigung von Messfehlern. Die wahren Parameter werden mit zunehmender Streuung der latenten Kovariablen immer stärker unterschätzt und die Volatilität der Schätzer steigt an. Insgesamt fällt der Effekt jedoch deutlich geringer aus, als im Messfehlerfall. Der Schätzfehler bei der Basisintensität $\lambda_{0(-1)}$ fällt zunächst positiv aus und entwickelt sich schließlich mit zunehmender Streuung der V_{hi} ins Negative. Eine mögliche Erklärung für dieses recht sonderbare Verhalten liegt in der Gestalt des Schätzers für λ_{0h} . Dieser ist gegeben durch

$$\hat{\lambda}_{0h} = \frac{\sum_{i=1}^n N_{hi}(\tau)}{\sum_{i=1}^n \int_0^\tau Y_{hi}(t) \exp\left(\hat{\boldsymbol{\beta}}_h^\top \mathbf{X}_i(t)\right) dt}. \quad (4.53)$$

Es ist denkbar, dass vernachlässigte Kovariablen insgesamt zu einer geringen Anzahl an beobachteten Sprüngen führen, als es beim angenommenen Modell der Fall sein müsste

und der Zähler von (4.53) somit zu niedrig ausfällt. Der Nenner von (4.53) wird allerdings aufgrund der negativen Verzerrung der Schätzer für die Kovariableneinflüsse ebenfalls unterschätzt. Es ist also anzunehmen, dass die Unterschätzung des Zählers von (4.53) bei relativ kleinen Varianzen der vernachlässigten Kovariablen geringer ausfällt als die des Nenners. Bei hoher Streuung verändert sich das Verhältnis allerdings zu Gunsten des Zählers. Trifft diese Erklärung zu, müsste sich der beobachtete Effekt mit zunehmender Anzahl an Kovariablen verstärken. Die Auswirkungen von vernachlässigten Kovariablen auf den Erwartungswert der Schätzer sind also immer noch deutlich erkennbar, fallen je-

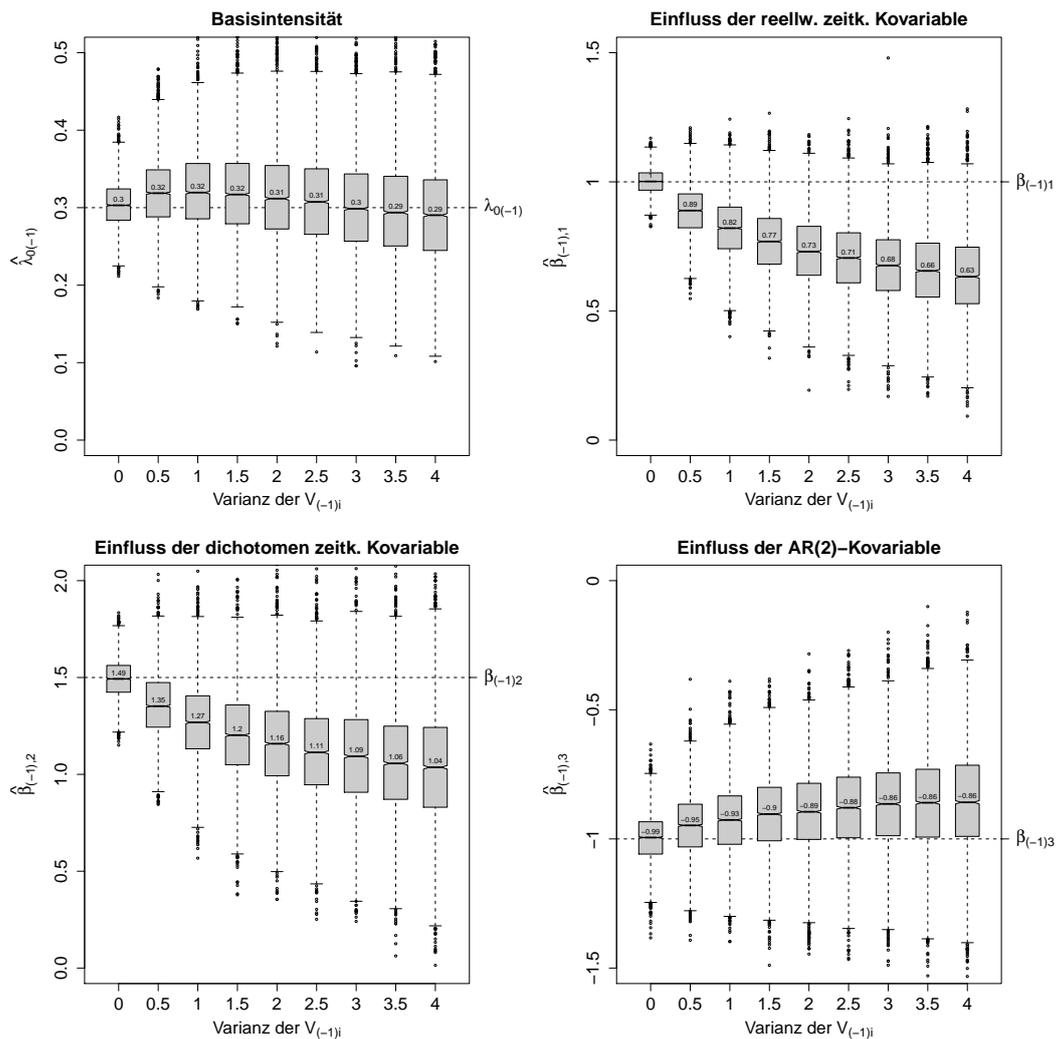


Abbildung 4.2: Lage und Streuung der Schätzer für die Parameter der Downgrade-Intensität $\lambda_{0(-1)}(t)$ bei Vernachlässigung von zeitkonstanten Kovariablen ($n = 250$)

Tabelle 4.2: Güte des Tests bei Vernachlässigung von zeitkonstanten Kovariablen ($\alpha = 5\%$)

Varianz	Stichprobengröße					
	25	50	100	250	500	1000
0.5	0.102	0.313	0.781	0.999	1.000	1.000
1.0	0.205	0.633	0.969	1.000	1.000	1.000
1.5	0.305	0.773	0.988	1.000	1.000	1.000
2.0	0.368	0.837	0.989	1.000	1.000	1.000
2.5	0.413	0.865	0.991	1.000	1.000	1.000

doch geringer aus als bei durch Messfehler verunreinigten Daten. Allerdings steigt die Unsicherheit der Schätzer mit zunehmender Streuung der V_{hi} weitaus schneller als in der Simulation aus Kapitel 3.2. Somit stellen auch vernachlässigte Kovariablen ein beträchtliches Problem bei der Risikobewertung seitens der Bank dar. Tabelle 4.2 zeigt, dass der Test ebenfalls sensibel auf diese Art der vernachlässigten Datenheterogenität reagiert. Die Güteeigenschaften scheinen dabei ähnlich zu sein wie bei der Vernachlässigung von Messfehlern. Dieses Ergebnis legt den Verdacht nahe, dass der Test sensibel auf jegliche Form von latenter Datenheterogenität reagiert, die zu einer Verzerrung der Parameterschätzer führt. Ob dies tatsächlich der Fall ist, bleibt im Detail zu untersuchen.

5 Empirische Analyse von Ratingwechselintensitäten und -wahrscheinlichkeiten

5.1 Daten

Bei dem vorliegenden Datensatz handelt es sich um eine interne Ratinghistorie aus dem Geschäftskundenportfolio einer großen deutschen Bank über den Zeitraum vom 27.08.1997 bis zum 05.09.2007. Das Ratingsystem der Bank umfasst 20 solvente Ratingklassen, wobei 1 das bestmögliche Rating darstellt. Die Insolvenz ist mit 21 kodiert. Der Datensatz beinhaltet insgesamt 760 Unternehmen, an denen 801 Ratingwechsel beobachtet wurden. Weitere deskriptive Statistiken zu dem Datensatz sind in Tabelle 5.1 zusammengefasst. Zusätzlich zu den Ratingbewegungen der Schuldner liegen Informationen zu Branche und Standort vor, wobei die Zuordnung nicht immer eindeutig ist, sondern auch mit Hilfe

Tabelle 5.1: Informationen zum Ratingdatensatz

Zensierte Historien	745					
Ausfälle durch Insolvenz	15					
beobachtete Ratingwechsel						
	Anzahl Ratingstufen					
	1	2	3	4	5	> 5
Downgrades	255	85	40	22	6	17
Upgrades	233	94	17	17	8	8
Gesamt	488	179	57	39	14	24
Anzahl Ratingwechsel je Schuldner						
	Anzahl Ratingwechsel					
	0	1	2	3	4	> 4
Anzahl Schuldner	371	174	111	60	23	21

von Gewichten erfolgen kann, wenn ein Schuldnerunternehmen Standorte in verschiedenen Ländern hat oder in mehreren Branchen tätig ist. Die Gewichte der Unternehmen in den einzelnen Branchen und Standorten werden bei der späteren Ratingmodellierung als Kovariablen verwendet.

Sämtliche Schuldnerunternehmen lassen sich - zumindest teilweise - den zehn Ländern Brasilien, China, Deutschland, Frankreich, Indien, Italien, Japan, Spanien, USA und Vereinigtes Königreich (UK) zuordnen. Zu diesen Ländern liegen jeweils drei makroökonomische Zeitreihen vor, welche als Kovariablen bei der Ratingmodellierung verwendet werden sollen: vierteljährliche Beobachtungen des Bruttoinlandsprodukts, und jeweils monatliche Beobachtungen der Industrieproduktion und der Einzelhandelsumsätze. Um daraus brauchbare Kovariablen zu erhalten, wird jeweils die Änderungsrate gegenüber dem Wert des Vorjahres betrachtet (im Folgenden durch Δ gekennzeichnet). Auf diese Weise werden linearer Trend und saisonale Effekte eliminiert. Beim BIP wird zusätzlich noch die durchschnittliche Steigung der Änderungsrate innerhalb des letzten Jahres ($\Delta^2\text{BIP}$) als weitere Kovariable berücksichtigt. Diese Steigung lässt sich als Anzeichen für die zukünftige Entwicklung der Wirtschaft in einem Land auffassen und stellt somit einen vermeintlichen Frühindikator für Veränderungen der makroökonomischen Situation dar.

Außerdem wird der Einfluss der monatlichen Durchschnittsrendite und -volatilität von Aktienindizes untersucht. Die Wahl des Aktienindex hängt dabei jeweils vom Unternehmensstandort ab und unterliegt folgendem Zuordnungsschema:

- Europa \rightarrow STOXX50
- USA \Rightarrow SP500
- Japan \Rightarrow NIKKEI
- Indien \Rightarrow BSE30
- China \Rightarrow Shenzhen Composite Index

Als siebte und letzte makroökonomische Kovariable werden 3-Monats-Zinssätze für das Interbankengeschäft betrachtet. Dazu liegen monatliche Beobachtungen des LIBOR (London Interbank Offered Rate) in den Währungen US-Dollar (USD), Britische Pfund (GBP) und Japanische Yen (JPY), des EURIBOR (Euro Interbank Offered Rate) und des SHIBOR (Shanghai Interbank Offered Rate) vor. Auch hier erfolgt die Zuordnung der Zeitreihen zu den Unternehmen gemäß des Standorts:

- Vereinigtes Königreich \rightarrow LIBOR (GBP)
- Europa (außer Vereinigtes Königreich) \rightarrow EURIBOR
- Japan \rightarrow LIBOR (JPY)

- Asien (außer Japan) → SHIBOR
- USA, sonstige Länder → LIBOR (USD)

Die Trennung von Japan und dem Rest Asiens wurde gewählt, da Japan im Gegensatz zum übrigen asiatischen Raum sehr stark industrialisiert ist und der JPY-LIBOR und der SHIBOR sich in ihrem Verhalten stark unterscheiden (innerhalb des Beobachtungszeitraums ergibt sich lediglich eine Korrelation von 0,09). Länder außerhalb Europas, Asiens und den USA wird der USD-LIBOR zugeordnet, da es sich beim US-Dollar um eine weltweit einflussreiche Währung handelt. Da der EURIBOR erst seit Anfang 1999 existiert, wird er vor dieser Zeit durch den GBP-LIBOR ersetzt. Dieser ist über den Zeitraum von Januar 1999 bis Dezember 2007 von den vorliegenden Zeitreihen am stärksten mit dem EURIBOR korreliert.

Sämtliche Zeitreihen werden vor der Ratingmodellierung standardisiert, so dass sich jeweils der Mittelwert 0 und die Varianz 1 ergibt. Auf diese Weise lassen sich die Parameterschätzer des Modells später besser interpretieren. Der Schätzer für den Effekt einer makroökonomischen Kennzahl gibt demnach an, inwieweit sich die Intensität eines Ratingwechsels verändert, wenn die zugehörige Kovariable einen Wert annimmt, der eine Standardabweichung über dem Durchschnittswert des Beobachtungszeitraums liegt. Unterschiede in der durchschnittlichen Wirtschaftslage einzelner Länder innerhalb des Beobachtungszeitraums lassen sich durch die Standorteffekte berücksichtigen. Weist ein Schuldnerunternehmen Standorte in mehreren Ländern auf, wird ein mit den Standortanteilen gewichtetes Mittel der entsprechenden Zeitreihen gebildet und als Kovariable verwendet. Liegt zu einem Standort keine Zeitreihe vor, wird diese als konstant Null angenommen, was aufgrund der Standardisierung der übrigen Zeitreihen als durchschnittliche Wirtschaftslage interpretierbar ist. Man beachte jedoch, dass jedes im Datensatz auftauchende Unternehmen zumindest teilweise einem Standort zugeordnet werden kann, zu dem Daten vorliegen. Somit nehmen die makroökonomischen Kovariablen bei keinem Schuldner über die gesamte Zeit den Wert Null an. Standortanteile, zu denen keine Wirtschaftszeitreihen vorliegen, haben demnach lediglich eine Abschwächung der Kovariablen zur Folge. Dies ergibt Sinn, da ein Unternehmen, das beispielsweise Standorte in Deutschland und den Cayman-Inseln hat weniger stark von der deutschen Wirtschaftslage beeinflusst wird, als ein Unternehmen, das ausschließlich in Deutschland ansässig ist.

5.2 Modellanpassung

In diesem Abschnitt soll nun das in Kapitel 2 vorgestellte Modell

$$\lambda_{hi}(t) = Y_{hi}(t)\lambda_{0h} \exp\left(\beta_h^\top \mathbf{X}_i(t)\right), \quad h \in \{+1, -1\} \quad (5.1)$$

an die vorliegenden Daten angepasst werden. Dazu gilt es, zunächst eine Vorüberlegung bezüglich der zeitabhängigen makroökonomischen Kovariablen $X_{il}(t)$, $l \in \{1, \dots, 7\}$, anzustellen. $X_{il}(t)$ könnte dem jüngsten zum Zeitpunkt t bekannten Wert des Wirtschaftsindikators entsprechen. Aber auch ein Wert, der weiter in der Vergangenheit zurückliegt, ist denkbar. Es stellt sich also die Frage, mit welcher Verzögerung die zeitabhängigen Kovariablen im Modell berücksichtigt werden sollen, da es sich sowohl um Früh-, als auch um Spätindikatoren für die Entwicklung der Schuldnerbonität handeln kann. Um dieser Frage nachzugehen, werden zunächst Modelle aufgestellt, in denen jeweils lediglich einer der sieben Wirtschaftsindikatoren als Kovariable verwendet wird. Der Einfluss dieser einen Kovariable wird dann für eine Verzögerung zwischen 1 und 12 Monaten mit Hilfe der partiellen Log-Likelihood (2.13) geschätzt. Anhand der Funktionswerte der Log-Likelihood lässt sich dann entscheiden, welche Verzögerung sich am besten für die entsprechende Kovariable eignet. Die Ergebnisse dieser Verzögerungs-Analyse sind in den Abbildungen 5.1 und 5.2 zusammengefasst. Zusätzlich zu den Werten der partiellen Log-Likelihood ist jeweils die geschätzte Richtung des Einflusses, also das Vorzeichen von β_{hl} , angegeben. Kovariablen, bei denen dieses Vorzeichen mit wachsender Verzögerung willkürlich schwankt, werden später nicht im Modell berücksichtigt um ein möglichst stabiles Ergebnis zu erzielen. Ein solches Verhalten der Schätzer legt die Vermutung nahe, dass die entsprechenden Wirtschaftsindikatoren ohnehin keinen nennenswerten Einfluss auf die Bonitätsschwankungen der Schuldner haben.

Nach Abbildung 5.1 erscheint bei der BIP-Änderung eine Verzögerung von einem Monat angemessen. Bei der Steigung der BIP-Änderung steigt die Log-Likelihood dagegen mit wachsender Verzögerung immer weiter an, so dass bei dieser Kovariablen jeweils der um 12 Monate zurückliegende Wert betrachtet wird. Dieses Resultat ist nicht überraschend, da die BIP-Änderung auf in der Vergangenheit erzielten Leistungen der Wirtschaft basiert und somit einen nachlaufenden Indikator darstellt. Die Steigung der BIP-Änderung vermittelt hingegen einen Eindruck über die zukünftige Entwicklung der Wirtschaft und kann daher als Frühindikator verstanden werden. Beim Einfluss der Kovariable „ Δ^2 BIP“ auf die Downgrade-Intensität ist bei einer Verzögerung von drei Monaten einmalig ein negatives Vorzeichen zu beobachten. Diese Instabilität des Parameterschätzers ist jedoch zu vernachlässigen, da sie lediglich in dem Verzögerungsbereich mit den niedrigeren Werten der partiellen Log-Likelihood auftritt. Die Downgrade-Intensität reagiert scheinbar etwas

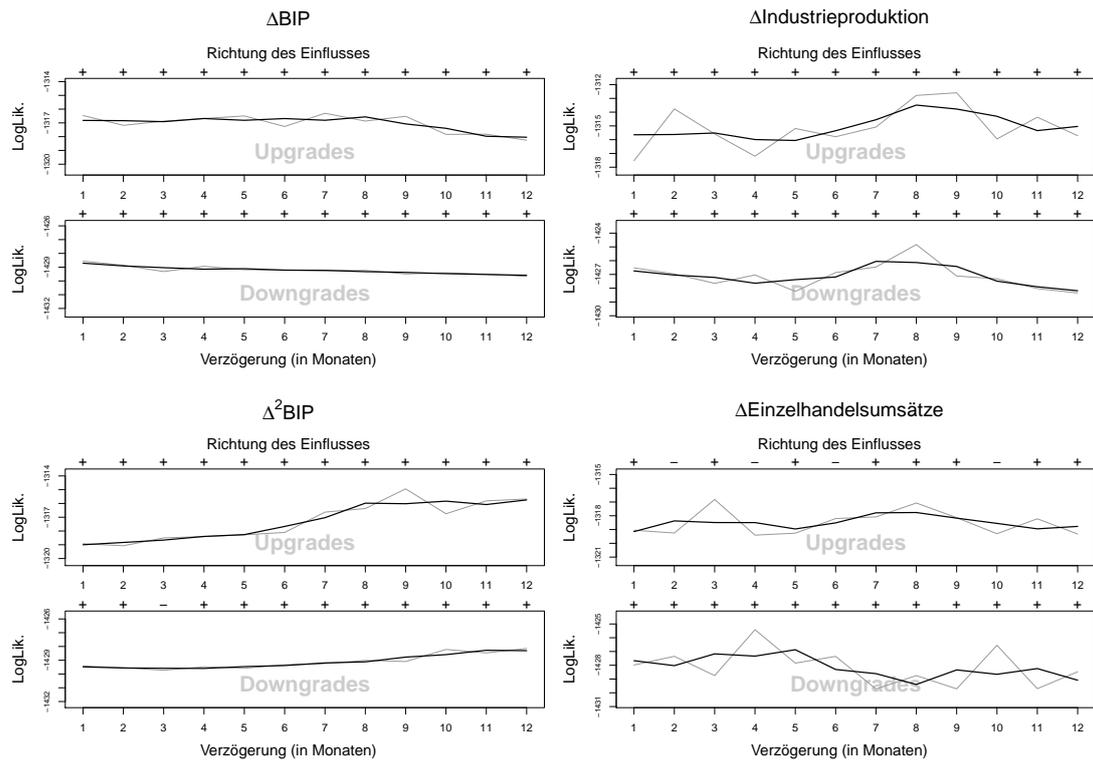


Abbildung 5.1: Verzögerungs-Analyse der Kovariablen „ Δ BIP“, „ Δ^2 BIP“, „ Δ Industrieproduktion“ und „ Δ Einzelhandelsumsätze“ (graue Linie = tatsächliche Werte der part. Log-Likelihood, schwarze Linie = gleitender Durchschnitt mit Fensterbreite 3)

schneller auf die Kovariable „ Δ Industrieproduktion“ als die Upgrade-Intensität. Hier fällt die Entscheidung zu Gunsten einer Verzögerung von 8 Monaten bei den Downgrades bzw. von 9 Monaten bei den Upgrades. Der Einfluss der Einzelhandelsumsätze auf die Upgrade-Intensität der Schuldner erweist sich als sehr instabil. Hier sind willkürliche Schwankungen des Vorzeichens von $\hat{\beta}_{+1l}$ zu beobachten, so dass diese Variable im Folgenden nicht bei der Modellierung der Upgrade-Intensität berücksichtigt wird. Der Einfluss der Änderung der Einzelhandelsumsätze auf die Downgrade-Intensität ist durchweg positiv. Da diese Kovariable bei der Upgrade-Intensität nicht berücksichtigt werden soll, hätte dies zur Folge, dass eine positive Entwicklung der Einzelhandelsumsätze zu einer Erhöhung der Downgradewahrscheinlichkeit führen würde. Dies ist aus ökonomischer Sicht nicht sinnvoll, so dass diese Kovariable komplett aus dem Modell gestrichen wird.

Die Ergebnisse der Verzögerungs-Analyse zu den übrigen drei Kovariablen sind in Abbildung 5.2 zusammengefasst. Demnach erscheint bei den Interbankzinssätzen eine Verzögerung

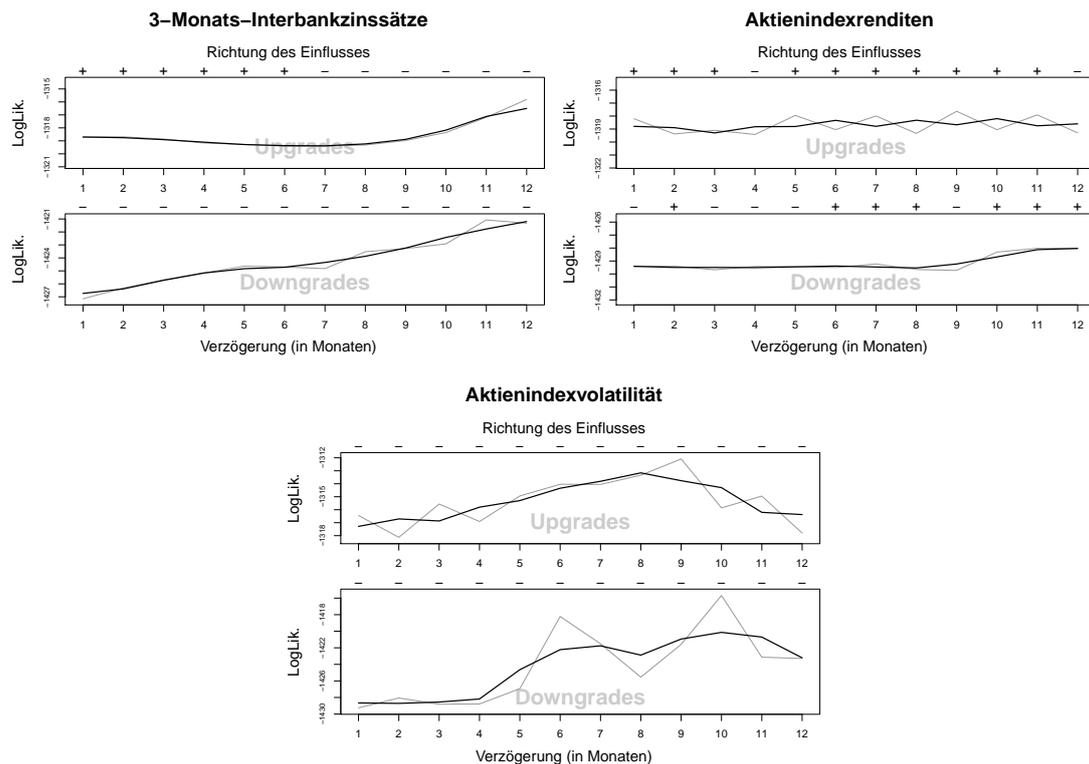


Abbildung 5.2: Verzögerungs-Analyse der Kovariablen „Interbankzinssätze“, „Aktienindexrendite“ und „Aktienindexvolatilität“ (graue Linie = tatsächliche Werte der part. Log-Likelihood, schwarze Linie = gleitender Durchschnitt mit Fensterbreite 3)

zung von 12 Monaten sinnvoll. Der Einfluss der Zinskovariablen auf die Upgrade-Intensität verändert sich mit wachsender Verzögerung vom Positiven ins Negative. Da es sich dabei jedoch nicht um eine willkürliche, sondern um eine systematische Schwankung handelt und der Schätzer für β_{+1l} im Bereich der höheren Werte der partiellen Log-Likelihood ein stabiles Vorzeichen aufweist, kann diese Kovariable bedenkenlos ins Modell aufgenommen werden. Bei den Aktienindexrenditen ist dies allerdings nicht der Fall. Hier schwankt die Richtung des Einflusses sowohl bei der Upgrade-, als auch bei der Downgrade-Intensität sehr stark. Aus diesem Grund wird diese Kovariable im späteren Modell keine Berücksichtigung finden. Der Effekt der Aktienindexvolatilität weist dagegen ein stabiles Vorzeichen auf. Die Downgrade-Intensität scheint dabei etwas langsamer auf diese Kovariable zu reagieren als die Upgrade-Intensität. In diesem Fall ist eine Verzögerung von 10 Monaten bei den Downgrades und 9 Monaten bei den Upgrades angemessen. Die gesamten Ergebnisse der Verzögerungs-Analyse sind in Tabelle 5.2 zusammengefasst.

Tabelle 5.2: Gewählte Verzögerungen der zeitabhängigen Kovariablen

	Verzögerung (in Monaten)	
	Upgrade-Modell	Downgrade-Modell
Δ BIP	1	1
Δ^2 BIP	12	12
Δ Industrieproduktion	9	8
Δ Einzelhandelsumsätze	-	-
Zins	12	12
Aktieindexrendite	-	-
Aktienindexvolatilität	9	10

Bevor das Modell geschätzt werden kann, gilt es auch noch, einige Vorüberlegungen bezüglich der Kovariablen „Branche“ und „Standort“ anzustellen. Die Daten beinhalten Unternehmen aus insgesamt 11 Standorten und 35 Branchen. Um eine vernünftige Schätzbarkeit der zugehörigen Effekte zu gewährleisten, werden jedoch nur Branchen und Standorte modelliert, zu denen sich mindestens 20 Schuldnerunternehmen zuordnen lassen. Schuldner aus einer zu dünn besetzten Kategorie werden zu der Branche bzw. dem Standort „Sonstige“ zusammengefasst. Somit bleiben 9 Standorte und 18 Branchen. Um die Interpretierbarkeit der Basisintensitäten zu gewährleisten, ist es außerdem nötig, jeweils eine Referenzkategorie zu wählen, deren Effekt nicht explizit modelliert wird. Die Basisintensität λ_{0h} entspricht der Intensität eines Ratingwechsels vom Typ h , wenn alle Kovariablen den Wert 0 annehmen. Dies ist genau dann der Fall, wenn sich ein Schuldner der Referenzbranche und dem Referenzstandort zuordnen lässt und die Wirtschaftskennzahlen eine durchschnittliche Ausprägung aufweisen. Ohne Referenzkategorien könnten dagegen niemals alle Kovariablen gleichzeitig den Wert 0 annehmen und die Basisintensität wäre nicht interpretierbar. Rund 57% der Schuldner aus dem Datensatz stammen aus Deutschland. Da dort also das Kerngeschäft der Bank liegt, wird dieser Standort als Referenz gewählt. Die Effekte der übrigen 8 Standorte sind demnach im Vergleich zu Deutschland zu interpretieren. Bei der Modellierung der Brancheneffekte wird ebenfalls die am stärksten besetzte Kategorie als Referenz gewählt. Dabei handelt es sich um die Branche „Sonstige“ mit rund 150 zugeordneten Schuldner. Die Wahl dieser Kategorie als Referenz ergibt Sinn, da es sich um eine Mischung aus einer Vielzahl verschiedener Branchen handelt. Somit ist davon auszugehen, dass Schuldner, die dieser Kategorie angehören, im Mittel ein durchschnittliches Ratingmigrationsverhalten aufweisen, was die Interpretation der 17 zu schätzenden Brancheneffekte begünstigt.

Nach diesen Vorüberlegungen kann das in Kapitel 2 vorgestellte Intensitätsmodell nun

angepasst werden. Die geschätzten Modellparameter sind in Tabelle 5.3 zusammengefasst. Der Einfluss der beiden BIP-Kovariablen ist intuitiv. Eine positive Entwicklung dieses Wirtschaftsindikators (Δ BIP), sowie ein positiver Ausblick in Zukunft (Δ^2 BIP) hat eine Verringerung der Downgrade-Intensität und eine Erhöhung der Upgrade-Intensität zur Folge. Negative Ausprägungen in beiden Variablen haben dementsprechend eine umgekehrte Wirkung. Im Gegensatz zu den übrigen makroökonomischen Kovariablen sind die geschätzten Effekte allerdings betragsmäßig recht klein. Bei der Entwicklung der Industrieproduktion würde man ein ähnliches Resultat erwarten. Allerdings weist die Kovariable „ Δ Industrieproduktion“ einen positiven Einfluss auf beide Intensitäten auf. Demnach kommt es in Zeiten, die durch eine überdurchschnittlich positive Entwicklung der Industrieproduktion geprägt sind zu erhöhten Ratingbewegungen in beide Richtungen. Die Tatsache, dass der Schätzer des Upgrade-Einflusses einen leicht höheren Wert annimmt, hat jedoch zur Folge, dass es im Mittel insgesamt zu einer leichten Ratingverbesserung kommt. Eine mögliche Erklärung für den auf den ersten Blick wenig intuitiven positiven Einfluss der Industrieproduktion auf die Downgrade-Intensität könnte in den Nachwirkungen einer Phase der erhöhten Produktion liegen. So könnte z.B. in Zeiten der Hochkonjunktur zu viel zusätzliches Personal eingestellt bzw. zusätzliche Maschinen angeschafft worden sein, was in wirtschaftlich ruhigeren Phasen eine zu große finanzielle Belastung darstellt. Denkbar sind auch Probleme beim Absatz der produzierten Güter. Hohe 3-Monats-Interbankzinssätze haben einen negativen Einfluss auf beide Intensitäten und somit einen stabilisierenden Einfluss auf die Ratings der Schuldner. Die verwendeten Referenzzinssätze sind ein Indikator für die generellen Konditionen am Geldmarkt. Im Umkehrschluss bedeuten die Parameterwerte also, dass es tendenziell zu starken Ratingbewegungen kommt, wenn 12 Monate zuvor Kredite einem günstigen Zins aufgenommen werden konnten. Günstige Konditionen am Geldmarkt verleiten Unternehmen zu Investitionen und Expandierungen, welche mittelfristig sowohl positive, als auch negative Folgen für die Bonität haben können. Hohe Volatilitäten haben Aktienmarkt haben ebenfalls eine stabilisierende Wirkung auf die Ratings der Schuldner, wobei sich die Downgrade-Intensität deutlich stärker verringert als die Upgrade-Intensität. Dies lässt den Rückschluss zu, dass die Bank die Ratings der Schuldner in wirtschaftlich unsicheren Zeiten tendenziell eher konstant lässt und insbesondere in Bezug auf Ratingherabstufung sehr zurückhaltend ist.

Wie bereits erwähnt, geben die Standorteffekte aufgrund der Standardisierung der makroökonomischen Kovariablen einen Aufschluss über die Wirtschaftslage in den Unternehmensstandorten im Vergleich zu Deutschland während der Kreditlaufzeit der dort ansässigen Schuldner. Auffällig ist hier insbesondere der Standort Japan. Hier scheint es zu verhältnismäßig wenigen Ratingbewegungen in beide Richtungen zu kommen. Bei

Tabelle 5.3: Geschätzte Modellparameter und Standardfehler

Wirtschaftskennzahlen	$\hat{\beta}_{+1t}$	$\hat{\sigma}(\hat{\beta}_{+1t})$	$\hat{\beta}_{-1t}$	$\hat{\sigma}(\hat{\beta}_{-1t})$
Δ BIP	0.051	(0.087)	-0.056	(0.081)
Δ^2 BIP	0.010	(0.108)	-0.064	(0.090)
Δ Industrieproduktion	0.259	(0.137)	0.232	(0.122)
Interbankzinssätze	-0.135	(0.097)	-0.132	(0.101)
Aktienindexvolatilität	-0.201	(0.122)	-0.422	(0.101)
Standorteffekte (Referenz: Deutschland)	$\hat{\beta}_{+1t}$	$\hat{\sigma}(\hat{\beta}_{+1t})$	$\hat{\beta}_{-1t}$	$\hat{\sigma}(\hat{\beta}_{-1t})$
USA	0.165	(0.255)	-0.127	(0.316)
UK	0.147	(0.371)	0.485	(0.222)
China	-0.039	(0.740)	0.783	(0.555)
Frankreich	0.542	(0.397)	-0.750	(0.587)
Spanien	0.272	(0.413)	0.623	(0.336)
Italien	-0.674	(0.725)	0.522	(0.324)
Japan	-1.396	(1.481)	-0.658	(0.911)
Sonstige	0.556	(0.330)	-0.270	(0.426)
Brancheneffekte (Referenz: Sonstige)	$\hat{\beta}_{+1t}$	$\hat{\sigma}(\hat{\beta}_{+1t})$	$\hat{\beta}_{-1t}$	$\hat{\sigma}(\hat{\beta}_{-1t})$
Betriebsmittel, Elektrizität & Müllbeseitigung	0.001	(0.232)	0.030	(0.225)
Öl & Bergbau	-0.185	(0.384)	-0.053	(0.337)
Stahl- & Metallprodukte	0.359	(0.266)	0.154	(0.302)
Transport	-0.004	(0.273)	0.513	(0.197)
Informationstechnik & Präzisionsinstrumente	0.305	(0.274)	-0.026	(0.363)
Chemikalien, Plastik & Gummi	-0.013	(0.343)	0.326	(0.362)
Automobil	0.680	(0.290)	0.469	(0.289)
Maschinen	0.021	(0.376)	-0.173	(0.416)
Bürobedarf (Großhandel)	-0.053	(0.526)	0.162	(0.359)
Dienstleister für Unternehmen	0.123	(0.300)	-0.208	(0.446)
Bau	-0.028	(0.573)	-0.173	(0.316)
Immobilien	0.232	(0.306)	0.145	(0.322)
Medizin & Pharma	-0.268	(0.530)	0.590	(0.385)
Konsumgüter (Produktion)	-0.391	(0.499)	-0.017	(0.359)
Lebensmittel, Getränke & Tabak	-0.291	(0.439)	0.352	(0.306)
Konsumgüter (Handel)	-0.191	(0.275)	0.477	(0.230)
Telefon, Rundfunk & Kabelfernsehen	-0.566	(0.471)	0.724	(0.304)
Basisintensitäten				
$\hat{\lambda}_{0(+1)} = 3.0 \cdot 10^{-4}$			$\hat{\lambda}_{0(-1)} = 2.8 \cdot 10^{-4}$	

den Brancheneffekten lassen sich zyklische und konjunkturabhängige Industriezweige erkennen. So weisen etwa die Automobilbranche und die Stahl und Metallindustrie erhöhte Intensitäten in beide Richtungen auf, während an Unternehmen aus dem Bereich „Öl und Bergbau“ im Vergleich zur Referenzkategorie weniger Ratingbewegungen zu beobachten sind. Auffällig sind zudem die Effekte der Branche „Telefon, Rundfunk & Kabelfernsehen“, welche auf einen negativen Trend der Ratings in dieser Branche schließen lassen. Dies lässt sich dadurch erklären, dass ein Großteil der Unternehmen des Datensatzes aus Deutschland stammt und dort das Telefonmonopol der Telekom innerhalb des Beobachtungszeitraums gefallen ist. Dies hatte zur Folge, dass der Markt kurzzeitig stark von alternativen Telefonanbietern überschwemmt war, von denen sich jedoch ein Großteil nicht halten konnte. Die geschätzten Basisintensitäten erscheinen auf den ersten Blick recht klein. Der Grund dafür liegt jedoch schlicht in der Tatsache, dass die Zeitskala in den Daten in Tagen angegeben ist und Ratingwechsel ein sehr seltenes Ereignis darstellen (vgl. Tabelle 5.1).

5.3 Bestimmung von Ratingwechselwahrscheinlichkeiten

Um den Einfluss, den die modellierten Kovariablen auf die Ratingbewegungen der Schuldner haben zu veranschaulichen, werden im Folgenden Ratingwechselwahrscheinlichkeiten basierend auf den geschätzten Intensitäten bestimmt. Insbesondere Ein-Jahres-Ratingwechselwahrscheinlichkeiten stellen später einen wichtigen Inputfaktor in Kreditportfolio-modellen wie etwa CreditMetrics (Gupton et al. (1997)) dar. Dazu werden die kumulierten Intensitäten in einer 21×21 Matrix $\hat{\Lambda}(t)$ zusammengefasst. Sei $\mathbf{x}(t)$ ein Pfad der modellierten Kovariablen für den die Ratingwechselwahrscheinlichkeiten bestimmt werden sollen. Für die Einträge von $\hat{\Lambda}(t)$ gilt

$$\hat{\Lambda}_{pq}(t) = \begin{cases} \int_0^t \hat{\lambda}_{0(+1)} \exp(\hat{\boldsymbol{\beta}}_{+1}^\top \mathbf{x}(u)) du & \text{falls } p - q = +1 \text{ und } p \neq 21 \\ \int_0^t \hat{\lambda}_{0(-1)} \exp(\hat{\boldsymbol{\beta}}_{-1}^\top \mathbf{x}(u)) du & \text{falls } p - q = -1 \\ - \sum_{q: q \neq p} \hat{\Lambda}_{pq}(t) & \text{falls } p = q \\ 0 & \text{sonst} \end{cases}. \quad (5.2)$$

Sei nun $[t_1, t_2]$ ein interessierendes Zeitintervall und $t_1 = t_0^* < t_1^* < \dots < t_m^* = t_2$ eine „möglichst feine“ Zerlegung. Dann ist ein Schätzer für die Migrationsmatrix \mathbf{P} für das Zeitintervall $[t_1, t_2]$ und den Kovariablenpfad $\mathbf{x}(t)$ gegeben durch

$$\hat{\mathbf{P}}(t_1, t_2; \mathbf{x}(t)) = \prod_{i=1}^m \left(\mathbf{I}_{21} + \hat{\Lambda}(t_i^*) - \hat{\Lambda}(t_{i-1}^*) \right). \quad (5.3)$$

Dieser Schätzer geht auf Aalen und Johansen (1978) zurück.

Tabelle 5.4: Einfluss der makroökonomischen Kovariablen auf die Ein-Jahres-Ratingwechselwahrscheinlichkeiten eines Schuldners aus der Ratingklasse 10

Wirtschaftsindikatoren					Wechselwahrscheinlichkeiten					MRV
Δ BIP	Δ^2 BIP	Δ Ind.	Zins	Akt.	8	9	10	11	12	
0	0	0	0	0	0.005	0.090	0.819	0.082	0.004	0.009
-1	-1	-1	0	0	0.003	0.068	0.850	0.076	0.003	-0.010
1	1	1	0	0	0.009	0.117	0.781	0.087	0.005	0.038
0	0	0	-1	0	0.006	0.100	0.797	0.091	0.005	0.010
0	0	0	1	0	0.004	0.080	0.839	0.074	0.003	0.008
0	0	0	0	-1	0.007	0.102	0.765	0.117	0.009	-0.020
0	0	0	0	1	0.003	0.077	0.861	0.057	0.002	0.024

Im Folgenden soll der Einfluss der Wirtschaftslage auf die Ratingdynamik mit Hilfe von Ratingwechselwahrscheinlichkeiten veranschaulicht werden. Dies geschieht am Beispiel eines Schuldners aus der Ratingklasse 10. Als Branche und Standort wird jeweils die Referenzkategorie gewählt. Es bietet sich an, die Kovariablen „ Δ BIP“, „ Δ^2 BIP“ und „ Δ Industrieproduktion“ gemeinsam zu untersuchen, da sich die Ausprägungen dieser Indikatoren in Bezug auf die Wirtschaftslage recht eindeutig interpretieren lassen. Bei den Interbankzinssätzen und der Aktienvolatilität ist dies nicht ohne weiteres der Fall, so dass es sinnvoll ist, diese beiden Kovariablen separat zu betrachten. Die Wechselwahrscheinlichkeiten werden jeweils für einen Zeithorizont von einem Jahr berechnet und es werden jeweils die Kovariablenausprägungen 0,-1 und 1 betrachtet. Dies entspricht aufgrund der Standardisierung der Variablen einem durchschnittlichen Wert und Ausprägungen, die eine Standardabweichung unter- bzw. oberhalb des Durchschnitts liegen. Die Ergebnisse sind in Tabelle 5.4 zusammengefasst. Zusätzlich zu den Ratingwechselwahrscheinlichkeiten ist dort die mittlere Ratingveränderung (MRV) eines Schuldners angegeben. Die MRV lässt sich mit Hilfe der Wechselwahrscheinlichkeiten bestimmen und gibt die erwartete Veränderung der Bonität in Ratingstufen an, wobei ein positives Vorzeichen für eine Verbesserung des Ratings steht. Die bestimmten Wechselwahrscheinlichkeiten verdeutlichen einmal mehr, dass Ratingwechsel recht seltene Ereignisse sind. In einer Wirtschaftslage, die in jeglicher Hinsicht durchschnittlich ist (alle Kovariablen nehmen den Wert 0 an), liegt die Wahrscheinlichkeit, dass sich das Rating eines Schuldners innerhalb eines Jahres nicht verändert bei rund 82%. Eine positive Wirtschaftssituation in Form von überdurchschnittlichen Ausprägungen der Kovariablen „ Δ BIP“, „ Δ^2 BIP“ und „ Δ Industrieproduktion“ führt nicht nur zu einer Verbesserung der mittleren Rating-

veränderung, sondern auch zu einer Erhöhung der Ratingwechselwahrscheinlichkeiten in beide Richtungen. Eine schlechte Wirtschaftslage hat dementsprechend zur Folge, dass es im Mittel zu einer Herabstufung der Schuldner kommt. Das Rating verhält sich jedoch insgesamt stabiler. Die Wahrscheinlichkeit, dass ein Schuldner in seiner Ratingklasse verbleibt, steigt in diesem Fall auf 85%. Wie sich bereits in der Intensitätsanalyse angedeutet hat, haben die 3-Monatsinterbankzinssätze nahezu keinen Einfluss auf die mittlere Veränderung des Ratings, sondern lediglich auf dessen Varianz. Hohe Volatilitäten am Aktienmarkt haben ebenfalls eine Verringerung der Ratingvariabilität zur Folge. Der Einfluss der Aktienindexvolatilität auf die erwartete Ratingveränderung ist aus ökonometrischer Sicht allerdings nicht ganz schlüssig, da eine hohe Stabilität am Aktienmarkt dazu führt, dass sich das Rating eines Schuldners im Mittel verschlechtert.

5.4 Testergebnisse im Ratingmodell

Der in Abschnitt 4.1 entwickelte Score-Test wird nun auf das in Abschnitt 5.2 geschätzte Modell angewendet. Für die Nullhypothese „ $H_0 : \eta_{+1} = \eta_{-1} = 0$ “ nimmt die χ^2_2 -Teststatistik (4.21) einen Wert von 7,61 an, was einen p-Wert von 0,022 zur Folge hat. Somit kann die Nullhypothese zu einem Niveau von 5% verworfen werden. Es besteht also Grund zur Annahme, dass die vorliegenden Daten Messfehler und/oder andere Formen der latenten Heterogenität aufweisen. In Anbetracht der Tatsache, dass eine Vielzahl von Kovariablen im Modell berücksichtigt wurde, sind Messfehler jedoch die wahrscheinlichere Ursache für das Testergebnis. Aufgrund der Resultate der Simulationsstudie in Abschnitt 3.2 ist demnach anzunehmen, dass die wahren Basisintensitäten und Kovariableneinflüsse in Kapitel 5.2 - und damit auch die Ratingwechselwahrscheinlichkeiten in Kapitel 5.3 - unterschätzt wurden.

6 Zusammenfassung und Fazit

Bei der Parametrisierung von Kreditportfoliomodellen spielen Ratingwechselwahrscheinlichkeiten eine wichtige Rolle. Die Schätzung dieser Wahrscheinlichkeiten erfordert die Modellierung der Ratingdynamik. Da sich das Rating eines Schuldners jederzeit ändern kann, ist dabei ein zeitstetiger Ansatz angemessen. Ein solches Modell setzt jedoch voraus, dass die zur Schätzung verwendeten Daten auf einer kontinuierlichen Überwachung der Ratings basieren. Nach Mählmann (2006) ist diese Voraussetzung in der Praxis jedoch nicht zwangsläufig erfüllt. Wird die Bonität der Schuldner lediglich sporadisch überprüft, kann es zu Messfehlern bei der Erfassung der Ratingwechselzeitpunkte kommen. Mit dieser Problematik beschäftigt sich die vorliegende Arbeit.

Bei dem Rating eines Schuldners handelt es sich um einen zeitstetigen Prozess mit endlichem Zustandsraum. Die Modellierung des Ratings über die Zeit erfolgt somit über die Intensitäten der Ratingwechsel. Für die Schätzung der Intensitätsparameter bietet sich ein zweistufiges Maximum-Likelihood-Verfahren an, bei dem zunächst nur die Einflüsse der im Modell berücksichtigten Kovariablen geschätzt werden. Die Ergebnisse lassen sich dann verwenden um die übrigen Intensitätsparameter zu bestimmen. Die Konsistenz dieses zweistufigen ML-Schätzers wird in der vorliegenden Arbeit mit Hilfe von Martingalgrenzwerttheorie nachgewiesen. In einer Simulationsstudie wird allerdings deutlich, dass die Vernachlässigung von Messfehlern zu einer starken Verzerrung der ML-Schätzer führen kann. Sowohl die Intensitäten, als auch durch Kovariablen verursachte Schwanken in den Intensitäten werden mitunter stark unterschätzt. Dies hat letztendlich eine Fehleinschätzung des Kreditrisikos zur Folge.

Messfehler lassen sich bei der Parameterschätzung berücksichtigen, indem die Intensitätsmodelle um einen geeigneten Frailty-Term, d.h. einen latenten stochastischen Prozess, erweitert werden. Die Parameterschätzung gelingt dann mit Hilfe der marginalen Likelihood, d.h. mit dem Erwartungswert der Likelihood bezüglich der Frailty-Terme. Dieses recht einfache Prinzip wird jedoch durch die Details erschwert, die es zu berücksichtigen gilt. So hängen die Schätzer für die Intensitätsparameter stark von der parametrischen Form der Verteilung ab, die den Messfehlern unterstellt wird. Aufgrund der Tatsache, dass es sich um latente Größen handelt, fällt es jedoch schwer, hier eine fundierte Entscheidung zu treffen. Um dieses Problem zu umgehen, schlagen Heckman und Singer (1984b) ein

nichtparametrisches Verfahren vor, welches jedoch recht instabil und rechenaufwendig ist.

Die Berücksichtigung von Messungenauigkeiten in den Ratingwechselzeitpunkten ist also keineswegs unproblematisch. Um sich diesem Aufwand nur im Bedarfsfall auszusetzen, ist es sinnvoll, zunächst mit Hilfe eines statistischen Tests zu untersuchen, ob die vorliegenden Daten Messfehler aufweisen. Ein solcher Test wird in der vorliegenden Arbeit nach einer Idee aus der Überlebenszeitanalyse (Lancaster (1990)) entwickelt. Von Interesse ist dabei, ob sich die Varianz der Messfehler signifikant von Null unterscheidet. Ist dies nicht der Fall liegen keine Messungenauigkeiten in den Ratingwechselzeitpunkten vor. Der Ansatz ist derselbe, wie bei der Berücksichtigung von Messfehlern: Basis für den Test ist der Erwartungswert der Likelihood eines Frailty-Intensitätsmodells. Um eine Verteilungsannahme an die Messfehlergrößen zu vermeiden, wird die Frailty-Likelihood allerdings vor der Erwartungswertbestimmung durch eine Approximation ersetzt. Diese Approximation stellt eine gute Annäherung der wahren Likelihood in der Nähe der Nullhypothese dar und hat den Vorteil, dass für die Erwartungswertbildung lediglich die ersten beiden Momente der Messfehlergrößen relevant sind. Ergebnis ist eine Likelihood, welche sowohl die interessierenden Intensitätsparameter als auch einen Messfehlervarianzparameter enthält und somit als Grundlage für das vorliegende Testproblem verwendet werden kann.

Aufgrund der Tatsache, dass der Varianzparameter nicht negativ sein kann, liegt er unter der Nullhypothese am Rand des zulässigen Parameterbereichs, was dazu führt, dass der ML-Schätzer nicht die üblichen Eigenschaften - wie etwa die asymptotische Normalverteilung - aufweist. Um einer Schätzung der Messfehlervarianz aus dem Weg zu gehen, bietet sich ein Score-Test auf Basis der erwarteten approximierten Likelihood an. Eine Schätzung des Varianzparameters ist hierbei nicht nötig, da dieser durch die Nullhypothese festgelegt wird. Ein Score-Test basiert auf der Annahme, dass der Score-Vektor - d.h. die erste Ableitung der (erwarteten approximierten) Log-Likelihood - an der Stelle der wahren Parameterwerte im Mittel den Wert Null annimmt. Dass diese Annahme auch im vorliegenden Fall zutrifft, folgt mit Hilfe einer Approximation der Dichte der Zwischen Eintrittszeiten der Ratingwechsel. Die gleichmäßige Konvergenz des Score-Vektors gegen eine normalverteilte Zufallsvariable lässt sich mit einem Satz von Donsker nachweisen. Zusammen mit der Konsistenz des zweistufigen ML-Schätzers ergibt sich eine χ^2 -verteilte Teststatistik für das vorliegende Testproblem.

Im Rahmen einer Simulationsstudie zeigen sich die Eigenschaften des Tests in endlichen Stichproben. Unter der Nullhypothese offenbart sich dabei ein leicht antikonservatives Verhalten. Das tatsächliche Testniveau konvergiert also von oben gegen den gewünschten Wert. In praxisrelevanten Stichprobengrößen ist dieser Effekt jedoch kaum noch zu erkennen. Der Test erweist sich bereits in moderaten Stichprobengrößen als recht trennscharf und erkennt Messfehler mit kleiner Varianz zuverlässig. Allerdings reagiert er nicht nur

sensibel auf Messfehler in den Ratingwechselzeitpunkten, sondern auch auf vernachlässigte Kovariablen. Diese haben allerdings die gleichen Konsequenzen für die Schätzer der Intensitätsparameter wie die Vernachlässigung von Messfehlern: Die wahren Werte der Parameter werden unterschätzt. Die Rückschlüsse, die sich aus einem signifikanten Testergebnis ziehen lassen bleiben demnach dieselben.

Für eine Anwendung des Tests stehen Ratinghistorien aus dem Geschäftskundenportfolio einer großen deutschen Bank zur Verfügung. Anhand dieser Daten wird ein Intensitätsmodell unter Berücksichtigung der Wirtschaftsindikatoren Bruttoinlandsprodukt, Industrieproduktion, Interbankzinssatz und Aktienvolatilität geschätzt. Als weitere Kovariablen dienen die Branche und der Standort der Schuldner. Einsetzen der geschätzten Parameter in die Teststatistik liefert einen signifikanten Wert. Es ist also davon auszugehen, dass die wahren Intensitätsparameter unterschätzt wurden. In Anbetracht der Tatsache, dass eine angemessene Anzahl an Kovariablen im Modell berücksichtigt wurde, sind Messfehler als wahrscheinliche Ursache für dieses Resultat anzusehen.

Messfehler in den Ratingwechselzeitpunkten scheinen demnach tatsächlich ein ernsthaftes Problem in der Praxis zu sein. Abhilfe schafft eine häufige und regelmäßige Kontrolle der Schuldnerratings, sowie eine Verbesserung des Informationsflusses zwischen Schuldner und Bank. Ist dies nicht in ausreichendem Maße möglich, müssen Messfehler bei der Bestimmung der Ratingwechselwahrscheinlichkeiten durch Frailty-Modelle berücksichtigt werden. Einen weiteren Lösungsansatz für die Messfehlerproblematik liefern Bladt und Sørensen (2009). Die in dieser Arbeit vorgestellte Methode erlaubt die Schätzung der Wechselintensitäten auf Grundlage von zeitdiskreten Beobachtungen. Somit wäre lediglich eine regelmäßige Kontrolle der Schuldnerratings nötig. Denkbar ist hier zum Beispiel ein Zeitabstand von jeweils einem Jahr. Die vorgestellte Methode setzt allerdings zeitkonstante Intensitäten voraus. Inwiefern sich die Resultate auf zeitabhängige Wechselraten unter Berücksichtigung von Kovariablen erweitern lassen ist fraglich.

Eine Vernachlässigung von Messungenauigkeiten führt zu einer Unterschätzung der Wechselwahrscheinlichkeiten und hat demnach zur Folge, dass die benötigte Eigenkapitalabsicherung für das Kreditrisiko der Bank zu niedrig eingeschätzt wird. Eine Verbesserung der Überwachungskontinuität oder die Berücksichtigung von Messfehlern geht also voraussichtlich mit einer Erhöhung der Kosten einher und ist somit nicht unbedingt im Interesse der Bank. Es liegt also insbesondere an der Bankenaufsicht auf diese Problematik hinzuweisen und die Banken diesbezüglich zu kontrollieren.

R-Code

R-Code zu den in der Arbeit durchgeführten Simulationen (ein Iterationsschritt):

```
score.me <- function(x){
# x[1]=Index zur Auswahl der Gamma-Messfehlervarianz (1->0, 2->0.25, 3->0.5,...)
# x[2]=Stichprobengröße

### SIMULATIONSPARAMETER #####

gam.j <- x[1] #Index der Messfehler-Varianz
N <- x[2] #Stichprobengröße

bi <- 0.3 #Basisintensität
betTU <- c(-1,1.5,1) #Kovariableneinflüsse Upgrades
betTD <- c(1,1.5,-1) #Kovariableneinflüsse Downgrades

EVZens <- 2.5 #Erwartungswert der Exp-Zensierung
MaxZens <- 10 #Maximaler Beobachtungszeitraum

### UNTERFUNKTIONEN #####

library(survival)
library(MASS)

#Fkt. zum Erzeugen von zuf. Verweildauern auf Basis von zeitabhängigen kov.
rtimev.v <- function(kovt, kov1, kov2, tend=3, ib=0.1, b=1, b1=1, b2=1, l0=1, v){
  #Wert, den das Integral annehmen muss
  hL <- -log(runif(1))/(10*exp(b1*kov1+b2*kov2)*v)
  #Werte die das Integral an den Sprungstellen der Kov annimmt
  if (t0==0){
    hR <- cumsum(exp(b*kovt)*ib)
  }else if (t0>0 & t0<tend-ib){
    hR <- cumsum(c(exp(b*kovt[floor(t0/ib+1)])*(1-(t0/ib-floor(t0/ib)))*ib,
    exp(b*kovt[(floor(t0/ib+1)+1):(tend/ib)])*ib))
  }else if (t0>=tend-ib & t0<tend){
    hR <- exp(b*kovt[floor(t0/ib+1)]*(t0/ib-floor(t0/ib))*ib
  }else{
    hR <- 999999999999
  }
  #Index des Intervalls in dem hL liegt
  hI <- cut(hL, breaks=c(0,hR,Inf), labels=F)
  if (hI > 1){
    #Länge des ersten Intervalls + Gesamtlänge der
    #vollständig durchlaufenen Zeitintervalle + Rest
    event <- (1-(t0/ib-floor(t0/ib))*ib+(hI-2)*ib+(hL-hR[hI-1])/
```

```

        (exp(b*kovt[min(c(length(kovt),floor(t0/ib)+hI)])))
    }else{
        #Integralwert / Funktionshöhe = Breite des
        #Integrationsintervalls (entspr. obigem Rest)
        event <- hL/(exp(b*kovt[floor(t0/ib)+hI]))
    }
    return(event)
}

# Simulation des Ratingprozesses
sim.rat.tv.v <- function(r0, Z, mZ, K, K1, K2, l0U, l0D, bU, b1U, b2U, bD, b1D, b2D,
    gsU, grU, gsD, grD){
    # Ausgangsratings, Zensierungszeiten, maximale Zensierungszeit, Kovariablendataframe,
    # Basisintensität Up- und Downgrades, Kovariablenparameter Up- und Downgrades,
    # Gamma-Term Parameter
    n <- length(Z)
    RH <- data.frame(ID=numeric(n), Start=numeric(n), Stop=numeric(n),
        Rating1=numeric(n), Rating2=numeric(n))
    z <- 0
    for (i in 1:n){
        rn <- r0[i]
        zi <- 0
        tges <- 0
        repeat{
            z <- z+1
            zi <- zi+1
            if (gsD!=0){
                vwd <- rtimev.v(kov=K[i,], kov1=K1[i], kov2=K2[i], t0=tges, tend=mZ, b=bD,
                    l0=l0D, b1=b1D, b2=b2D, v=rgamma(1, gsD ,grD))
            }else{
                vwd <- rtimev.v(kov=K[i,], kov1=K1[i], kov2=K2[i], t0=tges, tend=mZ, b=bD,
                    l0=l0D, b1=b1D, b2=b2D, v=1)
            }
            if (rn!=1){
                if (gsD!=0){
                    vwd[2] <- rtimev.v(kov=K[i,], kov1=K1[i], kov2=K2[i], t0=tges, tend=mZ,
                        b=bU, l0=l0U, b1=b1U, b2=b2U, v=rgamma(1, gsU ,grU))
                }else{
                    vwd[2] <- rtimev.v(kov=K[i,], kov1=K1[i], kov2=K2[i], t0=tges, tend=mZ,
                        b=bU, l0=l0U, b1=b1U, b2=b2U, v=1)
                }
            }
        }
        #Ratingänderung (-1=Upgrade, +1=Downgrade)
        dr <- as.numeric(which.min(vwd)==1)-as.numeric(which.min(vwd)==2)
        RH[z,1] <- i
        RH[z,2] <- tges
        RH[z,3] <- min(c(vwd+RH[z,2],Z[i]))
        RH[z,4] <- rn
        RH[z,5] <- rn+dr
        rn <- rn+dr
        tges <- RH[z,3]
    }
}

```

```

    if (rn==21){
      break
    }
    if (RH[z,3]==Z[i]){
      RH[z,5] <- rn-dr
      break
    }
  }
}
}
RH <- cbind(RH, K[RH[,1],])
colnames(RH)[-1:5] <- paste("Kov1t",0:(length(K[1,])-1),sep="")
return(RH)
}

# Extrahieren der für die Schätzung relevanten Informationen aus der Ratinghistorie
updownonly <- function(his, ratdif){
# Input: Ratinghistorie, interessierende Ratingdifferenz
# (1 = einstufigs Upgrade, -1 = einstufiges Downgrades
if (ratdif==1 & any(his[,4]==1)){
  #Zeilen mit den interessierenden Events
  event <- as.numeric((his$Rating1-his$Rating2)==ratdif)
  #erste Beob der Schuldner in der Ratingkl. 1
  idchange <- as.numeric(c(T,his[-1,1]-his[-length(his[,1]),1]!=0) & his[,4]==1)
  lastobs <- c(as.numeric(his[-1,1]-his[-length(his[,1]),1]!=0),1)
  #Zeilen mit der letzten Beobachtung eines Schuldners
  ar1 <- as.numeric(his[,4]==1)#Zeilen mit Ausgangsrating 1
  #Zeilen mit Event, IDchange oder letzte Beobachtung
  relZ <- apply(cbind(as.logical(event),as.logical(idchange),as.logical(lastobs),
    as.logical(ar1)), FUN=any, MARGIN=1)
  vwd <- cbind(his[which(relZ),c(1,2,3)], event[which(relZ)], his[which(relZ),4],
    his[which(relZ),-(1:5)]) #neuer Dataframe mit rel. Zeilen
  vwd[-1,2] <- vwd[-length(vwd[,1]),3] #Startzeiten korrigieren
  vwd[which(duplicated(vwd[,1])==F),2] <- 0 #Startzeiten korrigieren
  #Intervalle markieren, in denen kein Risiko besteht (Nullzeilen)
  vwd[which(vwd[,5]==1),2:4] <- 0
  vwd <- vwd[,-5] #Ausgangsratingspalte entfernen
  rownames(vwd) <- 1:length(vwd[,1])
  colnames(vwd) <- c("ID", "Start", "Stop", "Event", "Kov1")
  #Überflüssige Nullzeilen entfernen
  if (any(vwd$Stop==0)){
    #Nullzeilen entfernen, Ausnahme: Nullzeile ist einzige Beobachtung
    #eines Schuldners (technische Gründe)
    vwd <- vwd[order(vwd$Stop, decreasing=T),]
    if (any(vwd$Stop==0 & duplicated(vwd[,1]))){
      vwd <- vwd[-which(vwd$Stop==0 & duplicated(vwd[,1]) ),]
    }
    vwd <- vwd[order(vwd$ID, vwd$Start),]
    rownames(vwd) <- 1:length(vwd[,1])
  }
}
}else{
  event <- as.numeric((his$Rating1-his$Rating2)==ratdif)
  lastobs <- c(his[-1,1]-his[-length(his[,1]),1],1)

```

```

#hier ist idchange unnötig, da das Risikointervall stets bei 0 beginnt
relZ <- apply(cbind(as.logical(event),as.logical(lastobs)), FUN=any, MARGIN=1)
vwd <- cbind(his[which(relZ),c(1,2,3)], event[which(relZ)], his[which(relZ),-(1:5)])
vwd[-1,2] <- vwd[-length(vwd[,1]),3]
vwd[which(duplicated(vwd[,1])==F),2] <- 0
rownames(vwd) <- 1:length(vwd[,1])
colnames(vwd) <- c("ID", "Start", "Stop", "Event", "Kov1")
}
return(vwd)
}

#Fkt. zur Dataframe-Aufbereitung für die Parameterschätzung mit coxph()
coxtv.df <- function(dat, tend=3, ib=0.1){
#dat = Dataframe mit Spaltennamen ID, Start, Stop, Event und Kov1t0...
datC <- dat[1,1:5]
colnames(datC)[5] <- "Kov1"
z <- 1
for (h in 1:length(dat[,1])){
  if (dat$Stop[h]==0){
    datC[z,1] <- dat$ID[h]
    datC[z,-1] <- 0
    z <- z+1
  }else{
    #Anzahl der benötigten Einträge
    i1 <- cut(dat$Start[h], breaks=c(seq(0,tend-ib,ib),tend+10^(-6)), labels=F, right=F)
    i2 <- cut(dat$Stop[h], breaks=c(0,seq(ib,tend,ib)+10^(-6)), labels=F, right=F)
    L <- i2-i1+1
    if (L > 1){
      datC[z:(z+L-1),1] <- rep(dat[h,1],L)
      datC[z:(z+L-1),2] <- c(dat$Start[h],seq((floor(dat$Start[h]/ib)+1)
        *ib,floor((dat$Stop[h]-10^(-6))/ib)*ib,ib))
      datC[z:(z+L-1),3] <- c(seq((floor(dat$Start[h]/ib)+1)*ib,
        floor((dat$Stop[h]-10^(-6))/ib)*ib,ib),dat$Stop[h])
      datC[z:(z+L-1),4] <- c(rep(0,L-1),dat$Event[h])
      datC[z:(z+L-1),5] <- as.vector(dat[h,-(1:4)], mode="numeric")[i1:i2]
    }else{
      datC[z:(z+L-1),1] <- dat[h,1]
      datC[z:(z+L-1),2] <- dat$Start[h]
      datC[z:(z+L-1),3] <- dat$Stop[h]
      datC[z:(z+L-1),4] <- dat$Event[h]
      datC[z:(z+L-1),5] <- as.vector(dat[h,-(1:4)], mode="numeric")[i1]
    }
    z <- z+L
  }
}
return(datC)
}

#AR(2)-Prozess erzeugen
ar2 <- function(a1,a2,vari, n){
  x <- 0

```

```

x[2] <- 0
for(i in 3:(n+50)){
  x[i] <- a1*x[i-1]+a2*x[i-2]+rnorm(1,mean=0, sd=sqrt(vari))
}
return(x[-(1:50)])
}

#####

#mögliche Parameter für die Gamma-Vtlg. der Messfehler
gam.shapeU <- c(0,4,2,4/3,1,4/5,2/3,4/7,1/2,4/9,4/10,4/11,4/12,4/13,4/14,4/15,4/16)
gam.rateU <- c(0,4,2,4/3,1,4/5,2/3,4/7,1/2,4/9,4/10,4/11,4/12,4/13,4/14,4/15,4/16)
gam.shapeD <- c(0,4,2,4/3,1,4/5,2/3,4/7,1/2,4/9,4/10,4/11,4/12,4/13,4/14,4/15,4/16)
gam.rateD <- c(0,4,2,4/3,1,4/5,2/3,4/7,1/2,4/9,4/10,4/11,4/12,4/13,4/14,4/15,4/16)

#Auswahl der Gamma-Parameter
gam.shapeU <- gam.shapeU[gam.j]
gam.rateU <- gam.rateU[gam.j]
gam.shapeD <- gam.shapeD[gam.j]
gam.rateD <- gam.rateD[gam.j]

### DATENERZEUGUNG #####

#Kovariablen erzeugen
ko1 <- rnorm(N, mean=0, sd=sqrt(0.75))
ko2 <- sample(c(0,1), size=N, replace=T)
ko3 <- matrix(numeric(N*MaxZens*10), nrow=N)
for(k in 1:N){
  ko3[k,] <- ar2(a1=0.6, a2=0.2, n=MaxZens*10, vari=0.1)
}

#Ratingausgangsverteilung
V0 <- as.vector(rmultinom(1,size=N, prob=c(rep(3,5), rep(7,10), rep(3,5))))
#Ausgangsratings
ar0 <- rep( (1:20)[which(V0!=0)], V0[which(V0!=0)])
#Zensierungszeitpunkte
Zens <- pmin(rexp(N, 1/EVZens), rep(MaxZens,N))

#Ratinghistorie
rh <- sim.rat.tv.v(ar0, Z=Zens, mZ=MaxZens, K=ko3, K1=ko1, K2=ko2, l0U=bi, l0D=bi, bU=betTU[3],
  b1U=betTU[1], b2U=betTU[2], bD=betTD[3], b1D=betTD[1], b2D=betTD[2], gsU=gam.shapeU,
  grU=gam.rateU, gsD=gam.shapeD, grD=gam.rateD)

```

6 Zusammenfassung und Fazit

```
### PARAMETERSCHÄTZUNG #####

vwdU <- updownonly(rh, 1) #Aufbereitung der Historie in +1/-1 Sprünge
vwdD <- updownonly(rh, -1)

vwdUC <- coxtv.df(vwdU, tend=MaxZens) #weitere Aufbereitung für coxph()
vwdDC <- coxtv.df(vwdD, tend=MaxZens)

vwdUC <- cbind(vwdUC,ko1[vwdUC[,1]],ko2[vwdUC[,1]])
vwdUC <- vwdUC[,c(1:4,6:7,5)]
colnames(vwdUC)[5:7] <- c("Kov1","Kov2","Kov3")

vwdDC <- cbind(vwdDC,ko1[vwdDC[,1]],ko2[vwdDC[,1]])
vwdDC <- vwdDC[,c(1:4,6:7,5)]
colnames(vwdDC)[5:7] <- c("Kov1","Kov2","Kov3")

beU <- coxph(Surv(Start, Stop, Event) ~ Kov1+Kov2+Kov3+cluster(ID), data=vwdUC[which(vwdUC[,3]!=0),],
  robust=T, control=coxph.control(iter.max=100))$coefficient
beD <- coxph(Surv(Start, Stop, Event) ~ Kov1+Kov2+Kov3+cluster(ID), data=vwdDC,
  robust=T, control=coxph.control(iter.max=100))$coefficient
lOU <- sum(vwdUC$Event)/sum( exp(as.matrix(vwdUC[,5:7])%*%beU)*(vwdUC$Stop-vwdUC$Start) )
lOD <- sum(vwdDC$Event)/sum( exp(as.matrix(vwdDC[,5:7])%*%beD)*(vwdDC$Stop-vwdDC$Start) )

thetEU <- c(lOU,beU)
thetED <- c(lOD,beD)

### TESTSTATISTIK #####

#ID der Ereignisse
IDxU <- cumsum(apply(cbind(as.logical(c(1,vwdUC[-1,1]-vwdUC[-length(vwdUC[,1]),1])),
  as.logical(c(0,vwdUC[-length(vwdUC[,1]),4]))), FUN=any, MARGIN=1))
IDxD <- cumsum(apply(cbind(as.logical(c(1,vwdDC[-1,1]-vwdDC[-length(vwdDC[,1]),1])),
  as.logical(c(0,vwdDC[-length(vwdDC[,1]),4]))), FUN=any, MARGIN=1))
#kum. Intensität
ziU <- as.vector(by((vwdUC$Stop-vwdUC$Start)*thetEU[1]*exp(as.matrix(vwdUC[,5:7])%*%beU),
  INDICES=IDxU, FUN=sum))
ziD <- as.vector(by((vwdDC$Stop-vwdDC$Start)*thetED[1]*exp(as.matrix(vwdDC[,5:7])%*%beD),
  INDICES=IDxD, FUN=sum))
#Ableitung der kum. Intensität nach beta
ziU.db1 <- as.vector(by((vwdUC$Stop-vwdUC$Start)*thetEU[1]*vwdUC[,5]
  *exp(as.matrix(vwdUC[,5:7])%*%beU), INDICES=IDxU, FUN=sum))
ziD.db1 <- as.vector(by((vwdDC$Stop-vwdDC$Start)*thetED[1]*vwdDC[,5]
  *exp(as.matrix(vwdDC[,5:7])%*%beD), INDICES=IDxD, FUN=sum))
ziU.db2 <- as.vector(by((vwdUC$Stop-vwdUC$Start)*thetEU[1]*vwdUC[,6]
  *exp(as.matrix(vwdUC[,5:7])%*%beU), INDICES=IDxU, FUN=sum))
ziD.db2 <- as.vector(by((vwdDC$Stop-vwdDC$Start)*thetED[1]*vwdDC[,6]
  *exp(as.matrix(vwdDC[,5:7])%*%beD), INDICES=IDxD, FUN=sum))
ziU.db3 <- as.vector(by((vwdUC$Stop-vwdUC$Start)*thetEU[1]*vwdUC[,7]
  *exp(as.matrix(vwdUC[,5:7])%*%beU), INDICES=IDxU, FUN=sum))
ziD.db3 <- as.vector(by((vwdDC$Stop-vwdDC$Start)*thetED[1]*vwdDC[,7]
  *exp(as.matrix(vwdDC[,5:7])%*%beD), INDICES=IDxD, FUN=sum))
```

```

# Scorevektor
Scn <- matrix(numeric(N*10), nrow=N)
colnames(Scn) <- c("etaU", "lambOU", "betU1", "betU2", "betU3",
  "etaD", "lambOD", "betD1", "betD2", "betD3")

Scn[,1] <- as.vector(by(ziU^2-vwdU$Event*2*ziU, INDICES=vwdU[,1], FUN=sum))
Scn[,2] <- as.vector(by(vwdU$Event-ziU, INDICES=vwdU[,1], FUN=sum))/10U
coviu <- c((which(duplicated(IDxU)==F)-1)[-1],length(IDxU))
Scn[,3] <- as.vector(by(vwdU$Event*vwdUC[coviu,5]-ziU.db1, INDICES=vwdU[,1], FUN=sum))
Scn[,4] <- as.vector(by(vwdU$Event*vwdUC[coviu,6]-ziU.db2, INDICES=vwdU[,1], FUN=sum))
Scn[,5] <- as.vector(by(vwdU$Event*vwdUC[coviu,7]-ziU.db3, INDICES=vwdU[,1], FUN=sum))

Scn[,6] <- as.vector(by(ziD^2-vwdD$Event*2*ziD, INDICES=vwdD[,1], FUN=sum))
Scn[,7] <- as.vector(by(vwdD$Event-ziD, INDICES=vwdD[,1], FUN=sum))/10D
covid <- c((which(duplicated(IDxD)==F)-1)[-1],length(IDxD))
Scn[,8] <- as.vector(by(vwdD$Event*vwdDC[covid,5]-ziD.db1, INDICES=vwdD[,1], FUN=sum))
Scn[,9] <- as.vector(by(vwdD$Event*vwdDC[covid,6]-ziD.db2, INDICES=vwdD[,1], FUN=sum))
Scn[,10] <- as.vector(by(vwdD$Event*vwdDC[covid,7]-ziD.db3, INDICES=vwdD[,1], FUN=sum))

SC <- c(sum(Scn[,1]),sum(Scn[,6]))

# Informationsmatrix
IM <- matrix(numeric(100), nrow=10)
for (a in 1:10){
  for (b in 1:10){
    IM[a,b] <- sum( Scn[,a]*Scn[,b] )
  }
}

#Info-Matrix um lambda0 bereinigen -> Sigma-Matrix
Sig <- IM[-c(2,7),-c(2,7)]-IM[-c(2,7),c(2,7)]%*%ginv(IM[c(2,7),c(2,7)])%*%IM[c(2,7),-c(2,7)]

#Teststatistik
TS <- t(SC)%*%ginv(Sig[c(1,5),c(1,5)])%*%SC/N

### ERGEBNISSE #####
result <- c(TS, thetEU, thetED)
names(result) <- c("Teststat.", "10U", "betU1", "betU2", "betU3", "10D", "betD1", "betD2", "betD3")
return(result) #OUTPUT: Teststatistik und Parameterschätzer
}

```

Literaturverzeichnis

- O.O. Aalen und S. Johansen. An empirical transition matrix for nonhomogeneous Markov chains based on censored observations. *Scandinavian Journal of Statistics*, 5:141–150, 1978.
- P.K. Andersen, O. Borgan, R.D. Gill und N. Keiding. *Statistical Models based on counting processes*. Springer, New York, 1993.
- P.K. Andersen und R.D. Gill. Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, 10:1100–1120, 1982.
- Baseler Ausschuss für Bankenaufsicht. *Internationale Konvergenz der Kapitalmessung und Eigenkapitalanforderungen (Übersetzung der deutschen Bundesbank)*. Basel, 2004.
- P. Billingsley. *Statistical Inference for Markov Processes*. University of Chicago Press, Chicago, 1961.
- P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 1969.
- M. Bladt und M. Sørensen. Efficient estimation of transition rates between credit ratings from observations at discrete time points. *Quantitative Finance*, 9:147–160, 2009.
- A.A. Borovkov. *Mathematical Statistics*. Gordon & Breach, 1998.
- D.R. Cox. Regression models and life-tables. *Journal of the Royal Statistical Society, Ser. B*, 34:187–202, 1972.
- J. Davidson. *Stochastic Limit Theory*. Oxford, 1994.
- M. Fisz. *Wahrscheinlichkeitsrechnung und mathematische Statistik*. VEB Deutscher Verlag der Wissenschaften, Berlin, 1989.
- L.G. Godfrey. *Misspecification tests in econometrics*. Cambridge University Press, Cambridge, 1990.
- G.M. Gupton, C.C. Finger und M. Bhatia. Creditmetrics. Technical report, J.P.Morgan, 1997.

- J.J. Heckman und B. Singer. Econometric duration analysis. *Journal of Econometrics*, 24:63–132, 1984a.
- J.J. Heckman und B. Singer. A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52:271–320, 1984b.
- S. Jaggia. Alternative forms of the score test for heterogeneity in a censored exponential model. *The Review of Economics and Statistics*, 79:340–343, 1997.
- A. Kadam und P. Lenk. Bayesian inference for issuer heterogeneity in credit ratings migration. *Journal of Banking & Finance*, 32:2267–2274, 2008.
- J. Kim. Conditioning the transition matrix on wind. Technical report, RiskMetrics Group, 1999.
- Y.-D. Kim, L. James und R. Weißbach. Bayesian analysis of multi-state event history data: Beta-dirichlet process prior. *Biometrika*, 99:127–140, 2012.
- S.J. Koopman, R. Kräussl, A. Lucas und A.B. Monteiro. Credit cycles and macro fundamentals. *Journal of Empirical Finance*, 16:42–54, 2009.
- S.J. Koopman, A. Lucas und B. Schwaab. Modeling frailty-correlated defaults using many macroeconomic covariates. *Journal of Econometrics*, 162:312–325, 2011.
- T. Lancaster. *The Econometric Analysis of Transition Data*. Cambridge University Press, Cambridge, 1990.
- E. Lengart. Relation de domination entre deux processus. *Ann. Inst. Henri Poincaré*, 13:171–179, 1977.
- T. Mählmann. Estimation of rating class transition probabilities with incomplete data. *Journal of Banking & Finance*, 30:3235–3256, 2006.
- P. Nickell, W. Perraudin und S. Varotto. Stability of rating transitions. *Journal of Banking & Finance*, 24:203–227, 2000.
- G.G. Nielsen, R.D. Gill, P.K. Andersen und T.I.A. Sørensen. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics*, 19:25–43, 1992.
- C.R. Rao. Score test: Historical review and recent developments. In *Advances in Ranking and Selection, Multiple Comparisons, and Reliability*, pages 3–20. Birkhäuser, Boston, 2005.

- J. Trussel und T. Richards. Correcting for unmeasured heterogeneity in hazard models using the Heckman-Singer strategy. *Sociological Methodology*, 15:242–276, 1985.
- A.W. van der Vaart. *Asymptotic Statistics*. Cambridge University Press, 1998.
- R. Weißbach und T. Mollenhauer. Modelling rating transitions. *Journal of the Korean Statistical Society*, 4:469–485, 2011.
- R. Weißbach und R. Walter. A likelihood ratio test for stationarity of rating transitions. *Journal of Econometrics*, 155:188–194, 2010.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Dortmund, den 13. Juli 2012