

Quantitative Validierung von Ratingmodellen für Low Default Portfolios mittels Benchmarking

Die neuen Eigenkapitalbestimmungen von Basel II stellen eines der am heftigsten diskutierten Themen der Finanzbranche in der jüngeren Vergangenheit dar. Nach einem mehrjährigen Konsultationsprozess sind die Regelungen grundsätzlich per 1. Jänner 2007 in Kraft getreten – die fortgeschrittenen Ansätze (fortgeschrittener IRB-Ansatz und AMA-Ansätze) folgen per 1. Jänner 2008. Im Bereich des Kreditrisikos ermöglichen die neuen Bestimmungen die Ermittlung der Bemessungsgrundlage zur Berechnung des Eigenmittelerfordernisses mittels interner Ratingsysteme. Dies setzt eine Bewilligung vonseiten der Aufsichtsbehörden voraus, zu deren Erlangung die internen Ratingsysteme bestimmte Anforderungen erfüllen müssen. Eine dieser Anforderungen sieht vor, dass die verwendeten Modelle eine gute Vorhersagekraft besitzen, was einmal jährlich durch die Bank im Rahmen einer qualitativen und quantitativen Validierung der Modelle zu überprüfen ist. Die bei der quantitativen Validierung verwendeten statistischen Methoden benötigen zur Ableitung von validen Aussagen eine gewisse Datengrundlage, die bei Ratingmodellen für sogenannte Low Default Portfolios (LDP), also Portfolios, bei denen es nur sehr selten zu Ausfällen kommt, typischerweise nicht gegeben ist. Der vorliegende Beitrag befasst sich zunächst allgemein mit den Problemen von LDP im IRB-Ansatz und geht auf die Problematik der Validierung von Ratingmodellen für LDP ein. Im Anschluss daran wird eine alternative Vorgehensweise zur quantitativen Validierung solcher Modelle – basierend auf der Idee des Benchmarkings – dargestellt. Die Anwendung der vorgeschlagenen Validierungsmethode wird schließlich anhand eines Beispiels aufgezeigt.

Markus Ricke,
Georg von Pfölstl

1 Einleitung

Nach einem mehrjährigen Konsultationsprozess veröffentlichte der Basler Ausschuss für Bankenaufsicht (BAB) im Juni 2004 die überarbeitete Rahmenvereinbarung „Internationale Konvergenz der Eigenkapitalmessung und Eigenkapitalanforderungen“ (Basel II). Mit den neu gefassten EU-Richtlinien 2006/48/EG (Bankrechtsrichtlinie, CRD) und 2006/49/EG (Kapitaladäquanzrichtlinie, CAD) wurden die Regelungen von Basel II in das EU-Recht übernommen. Die nationale Umsetzung dieser Richtlinien erfolgte in Österreich durch die Überarbeitung des Bankwesengesetzes (BWG) im August 2006 und die Veröffentlichung der neuen Solvabilitäts- und Offenlegungsverordnung (SolvaV und Off-VO) im Oktober

2006. Mit dem 1. Jänner 2007 sind die neuen Eigenkapitalregelungen von Basel II in Österreich schließlich in Kraft getreten.¹

Im Bereich des Kreditrisikos ermöglichen die neuen Regelungen die Ermittlung der Bemessungsgrundlage zur Berechnung des Eigenmittelerfordernisses mittels interner Ratingsysteme (IRB-Ansatz gemäß § 22b BWG). Hierzu benötigen Banken allerdings eine Bewilligung vonseiten der Aufsichtsbehörden, zu deren Erlangung die internen Ratingsysteme eine Reihe von Anforderungen zu erfüllen haben, die in § 37 ff. SolvaV definiert sind.

Eine dieser Anforderungen besagt, dass die verwendeten Ratingmodelle nachweislich eine gute Vorhersagekraft besitzen müssen, was

Wissenschaftliche
Begutachtung:
Gerhard Winkler,
OeNB

¹ Durch die Ausübung eines nationalen Wahlrechts kann die Anwendung der neuen Regelungen vonseiten der österreichischen Kreditinstitute auf den 1. Jänner 2008 verschoben werden.

jedenfalls einmal jährlich durch die Bank im Rahmen einer qualitativen und quantitativen Validierung der Modelle zu überprüfen ist (§§ 41 und 59 SolvaV). Die bei der quantitativen Validierung typischerweise verwendeten statistischen Methoden setzen zur Ableitung von validen Aussagen eine gewisse Datengrundlage voraus, was bei Ratingmodellen für Low Default Portfolios (LDP), also Portfolios, bei denen es nur sehr selten zu Ausfällen kommt (z. B. Staaten), zu Problemen führt.

Das Ziel des vorliegenden Beitrags ist es daher, eine alternative Vorgehensweise zur quantitativen Validierung von Ratingmodellen aufzuzeigen, die zur Beurteilung der Vorhersagekraft von Ratingmodellen für typische LDP wie Staaten oder Banken angewendet werden kann. Die dargestellte Methode orientiert sich hierbei an der in Hornik et al. (2006) verwendeten Methode und basiert auf der Idee des Benchmarkings, also des Vergleichs der Ergebnisse des bankinternen Ratingmodells mit den Ergebnissen anderer Verfahren bzw. mit externen Daten. Der vorliegende Beitrag behandelt den Vergleich mit externen Daten.

Der Beitrag befasst sich zunächst mit den Problemen von LDP im IRB-Ansatz (Kapitel 2). In Kapitel 3 wird die Problematik bei der quantitativen Validierung von Ratingmodellen für LDP erläutert, bevor in Kapitel 4 eine alternative Vorgehensweise zur quantitativen Validierung von Ratingmodellen für LDP dargestellt wird, die auf der Idee des Benchmarkings beruht. In Kapitel 5 zeigen wir schließlich die Anwendung der vor-

geschlagenen Validierungsmethode anhand eines konkreten Beispiels. Der Beitrag endet mit einem Fazit in Kapitel 6.

2 Low Default Portfolios im IRB-Ansatz

Unter einem LDP werden solche Portfolios verstanden, bei denen keine oder nur sehr wenige Ausfalldaten vorliegen. Ein LDP lässt sich auf verschiedene Ursachen zurückführen, wie z. B.:²

- Portfolios mit wenigen Kunden – entweder generell (z. B. Staaten) oder auf der Ebene des Kreditinstituts
- eine generell niedrige Ausfallrate bei bestimmten Kundengruppen, (z. B. Banken)
- eine geringe Ausfallrate bei bestimmten Kundengruppen während bestimmter Zeitperioden
- eine kurze Ausfallhistorie aufgrund der Tatsache, dass das Kreditinstitut erst seit kurzer Zeit in einem bestimmten Portfoliosegment aktiv ist

Auf Basis dieser unterschiedlichen Ursachen wird häufig eine Unterteilung in die folgenden Formen von LDP vorgenommen:³

- *Langfristig vs. kurzfristig*: Langfristige LDP lassen sich auf generell niedrige Ausfallraten bestimmter Kundengruppen oder eine geringe Zahl an Kunden zurückführen; ein kurzfristiges LDP liegt hingegen dann vor, wenn der Grund für den geringen Umfang an Ausfalldaten darin liegt, dass ein Institut erst seit kurzem in einem neuen Marktsegment tätig ist.

² Siehe BBA und ISDA (2005).

³ Siehe CEBS (2006, S. 101).

- *Systematisch vs. institutsspezifisch:* Während bei systematischen LDP sämtliche Kreditinstitute mit der Problematik keiner oder nur sehr weniger Ausfalldaten konfrontiert sind, ist bei einem institutsspezifischen LDP nur das entsprechende Kreditinstitut von der Datenproblematik betroffen.

Obwohl der Mangel an Ausfalldaten bei LDP zu Problemen bei der Entwicklung und Validierung von Ratingmodellen sowie der Schätzung und Validierung von Risikoparametern für diese Portfolios führt, sind für LDP in den gesetzlichen Bestimmungen keine eigenen Anforderungen vorgesehen. Dies führte in vielen Kreditinstituten zu der Befürchtung, dass bei LDP die Anwendung eines IRB-Ansatzes grundsätzlich ausgeschlossen sei. Der BAB hat als Reaktion auf diesbezügliche Anfragen im September 2005 einen Newsletter veröffentlicht.⁴ Die Grundaussage des BAB ist, dass die Datenproblematik LDP nicht automatisch vom IRB-Ansatz ausschließe. Vielmehr sollten bei LDP alle Möglichkeiten zur Erweiterung der Datenbasis, sowohl mittels interner als auch externer Daten, ausgeschöpft werden, und in dem Fall, dass die Datenbasis anschließend immer noch nicht ausreichend groß wäre, sollten alternative Schätz- und Validierungsverfahren eingesetzt werden.⁵ Zudem hätten Kreditinstitute bei einer weniger zufriedenstellenden Datenbasis und einer dadurch bedingten größe-

ren erwarteten Fehlerspannweite bei der Parameterschätzung die Sicherheitszuschläge bei den Risikoparametern zu erhöhen.⁶

Beim Rating geht es um die Ordnung von Schuldern nach ihrer Bonität. Hierzu wird typischerweise eine diskrete Skala mit verschiedenen Ratingklassen verwendet. Zur Entwicklung eines Ratingmodells wird oftmals auf statistische Verfahren, z. B. die logistische Regression, zurückgegriffen, die allerdings eine gewisse Ausfalldatenbasis benötigen. Bei LDP können diese Verfahren somit nicht angewendet werden, sodass bei diesen Portfolios typischerweise Expertenmodelle, dies sind Modelle, bei denen die Auswahl und Gewichtung der verschiedenen Ratingkriterien durch Experten vorgenommen werden, zum Einsatz kommen.

Die Verwendung eines Expertenmodells für LDP ist grundsätzlich zulässig.⁷ Die Tatsache, dass es sich um ein Expertenmodell handelt, entbindet die Kreditinstitute jedoch nicht von der Verpflichtung, das Modell regelmäßig mittels quantitativer Verfahren zu validieren. Dies stellt viele Kreditinstitute vor Probleme, da die klassischerweise zur quantitativen Validierung eines Ratingmodells eingesetzten Verfahren eine gewisse Zahl an Ausfällen voraussetzen, die bei LDP nicht gegeben ist. Im folgenden Kapitel wird daher eine alternative Vorgehensweise zur quantitativen Validierung von Ratingmodellen für LPD dargestellt.

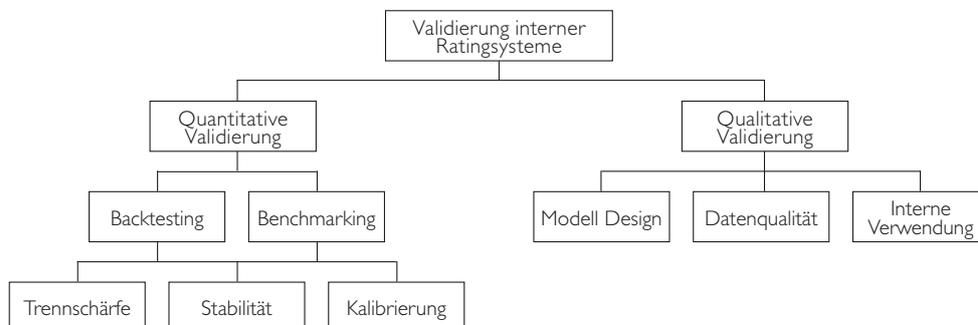
⁴ Siehe BCBS (2005a).

⁵ Diese Aussagen finden sich auch in dem im April 2006 veröffentlichten CP 10 des Committee of European Banking Supervisors (CEBS), siehe CEBS (2006).

⁶ In der nationalen Gesetzgebung findet sich dieser Punkt in § 47 Abs. 6 SolvaV.

⁷ Für die Parameterschätzung ist eine solche Vorgehensweise hingegen nicht zulässig, da in § 47 Abs. 1 SolvaV explizit gefordert wird, dass Parameterschätzungen nicht ausschließlich auf wertenden Annahmen basieren, sondern auch auf empirischen Ergebnissen beruhen.

Validierung eines Ratingmodells



Quelle: Deutsche Bundesbank (2003, S. 62).

3 Validierung von Low Default Portfolios

Gemäß Deutscher Bundesbank (2003) sowie OeNB und FMA (2004) hat die Validierung eines Ratingmodells die in Grafik 1 dargestellten Maßnahmen zu umfassen.

Während im Rahmen der quantitativen Validierung die Trennschärfe⁸ und die Korrektheit der Kalibrierung⁹ des Ratingmodells sowie die Stabilität der Ratingergebnisse mittels statistischer Verfahren untersucht werden, bezieht sich die qualitative Validierung auf die Datenqualität, das Modell Design und die interne Verwendung der Ratingergebnisse im Risikomanagement der Bank. Die quantitative Validierung kann hierbei auf Basis interner Daten (Backtesting) oder externer Daten (Benchmarking) durchgeführt werden.

Bei Ratingmodellen für LDP ist eine quantitative Validierung mittels Backtesting nur sehr eingeschränkt möglich, da die Zahl der Ausfälle in dem Portfolio der Bank typischerweise so gering ist, dass die Durchführung statistischer Tests nicht zu

sinnvollen Ergebnissen führt. Bei Vorliegen eines institutsspezifischen LDP kann diese Problematik durch die Verwendung von Daten anderer Kreditinstitute gelöst werden. Handelt es sich jedoch um ein systemisches LDP, ist auch eine quantitative Validierung mittels Benchmarking (im engen Sinn) nicht möglich.

In BCBS (2005a) sowie in BBA und ISDA (2005) wird der Begriff des Benchmarkings daher weiter definiert und umfasst hier auch Maßnahmen wie den Vergleich der bankinternen Ratings mit den Bonitätseinstufungen durch Ratingagenturen und mit aus Kapitalmarktpreisen abgeleiteten Proxies für das Ausfallrisiko. In dem folgenden Kapitel zeigen wir eine mögliche Vorgehensweise für einen solchen Vergleich.

4 Benchmarking von Ratingmodellen für systemische Low Default Portfolios

Bei einem Rating handelt es sich um eine ordinale Variable, d. h., die Schuldner werden nach ihrer Bonität gereiht, wobei typischerweise eine

⁸ Unter der Trennschärfe eines Ratingmodells wird die Fähigkeit verstanden, Schuldner, die ausfallen werden, von solchen, die nicht ausfallen werden, zu unterscheiden.

⁹ Unter der Kalibrierung eines Ratingmodells wird die Zuordnung von Ausfallwahrscheinlichkeiten zu den verschiedenen Ratingklassen verstanden.

diskrete Skala mit verschiedenen Ratingklassen verwendet wird, denen die Schuldner zugeordnet werden.¹⁰

Wir präsentieren im Folgenden einen Ansatz zur quantitativen Validierung von Ratingmodellen für systemische LDP mittels Benchmarking,¹¹ bei dem die ordinale Struktur der Ergebnisse eines Ratingmodells, also die Reihung der Schuldner nach ihrer Bonität, mit der von Ratingagenturen bzw. auf dem Kapitalmarkt beobachtbarer Proxies für das Ausfallrisiko verglichen wird.¹²

Als Methode zur Messung der Stärke und Richtung der Assoziation zwischen zwei ordinalskalierten Variablen werden in der Literatur oftmals Maße wie Spearman's Korrelationskoeffizient, Somer's D oder Kendall's τ vorgeschlagen.¹³ Emond und Mason (2002) haben jedoch gezeigt, dass diese Maße gewisse Schwächen aufweisen und daher ein verbessertes Maß, τ_x , vorgeschlagen, das beispielsweise in Hornik et al. (2006) zur Anwendung kommt.

Zur Berechnung von τ_x wird bei der Betrachtung eines Samples mit n Kreditnehmern zunächst für jedes

Merkmal¹⁴ eine $n \times n$ -Matrix erstellt, deren Elemente bei dem Merkmal a wie folgt bestimmt werden:

$a_{xy} = 1$, wenn Kreditnehmer x eine bessere oder die gleiche Ausprägung bei dem Merkmal aufweist wie Kreditnehmer y ,

$a_{xy} = -1$, wenn Kreditnehmer x eine schlechtere Ausprägung bei dem Merkmal aufweist als Kreditnehmer y , und

$a_{xy} = 0$ für alle Diagonalelemente der Matrix.

Mithilfe dieser Matrix kann anschließend das τ_x für die Merkmale a und b mit folgender Formel berechnet werden:

$$\tau_x = \frac{\sum_{x=1}^n \sum_{y=1}^n a_{xy} b_{xy}}{n(n-1)} \quad (1)$$

τ_x kann Werte zwischen -1 und $+1$ annehmen, wobei der Grad der Übereinstimmung umso höher ist, je höher τ_x ist.

Als Benchmark werden typischerweise externe Ratings sowie auf dem Kapitalmarkt beobachtbare Proxies für das Ausfallrisiko, z. B. Bondspreads¹⁵ oder Credit Default Swap

¹⁰ In einem weiteren Schritt, der Kalibrierung, wird den einzelnen Ratingklassen eine Ausfallwahrscheinlichkeit zugewiesen.

¹¹ Es sei an dieser Stelle erwähnt, dass für eine Benchmarking-Analyse zum einen ausreichend große Multi-Rater-Panels erforderlich sind. Zum anderen haben diese Panels vollständig zu sein, was in der Praxis in der Regel nicht gegeben ist, da nicht alle Agenturen alle betreffenden Kreditnehmer bewerten. Diese Problematik wird im vorliegenden Beitrag insofern ausgeklammert, als im Beispiel alle Kreditnehmer von allen Agenturen beurteilt werden. Zur angesprochenen Problematik siehe Hornik et al. (2006).

¹² Der in diesem Beitrag dargestellte Ansatz fällt somit unter den Bereich der Untersuchung der Trennschärfe; unter der Annahme, dass die Ratingagenturen bzw. Kapitalmarktteilnehmer in der Lage sind, die in Zukunft ausfallenden Schuldner von den nicht ausfallenden Schuldner zu trennen. Eine Überprüfung der Kalibrierung von Ratingmodellen für LDP mittels Benchmarking ist ebenfalls möglich, aber nicht Gegenstand dieses Beitrags.

¹³ Siehe dazu BCBS (2005b).

¹⁴ Merkmale sind hier das interne Rating sowie die für den Vergleich herangezogenen Proxies für die Ausfallwahrscheinlichkeit, z. B. ein externes Rating.

¹⁵ Ein Bondspread bezeichnet die Differenz aus der Rendite eines Bonds und der Rendite eines (nahezu) risikolosen Bonds mit derselben Laufzeit. Er ist typischerweise umso höher, je höher das Ausfallrisiko des Emittenten des Bonds ist.

(CDS)-Spreads,¹⁶ empfohlen.¹⁷ Hierbei wird implizit davon ausgegangen, dass die Reihung der Schuldner durch die externen Ratingagenturen bzw. durch die Kapitalmarktakteure perfekt ist.

Die Ratings der großen Agenturen und die Höhe von Bondspreads bzw. CDS-Spreads stehen in einem engen Zusammenhang.¹⁸ Dennoch gibt es einige wichtige Unterschiede zwischen diesen Maßen für das Ausfallrisiko eines Schuldners. Einer dieser Unterschiede betrifft die Stabilität des Maßes.

Ratingagenturen betonen, dass ihre Ratings „through-the-cycle“ (TTC) seien.¹⁹ Dies bedeutet, dass das Rating unabhängig vom Konjunkturzustand die langfristige Bonität des Schuldners abbilden soll.²⁰ Kurzfristige, eventuell nur temporäre Änderungen des Ausfallrisikos werden nicht abgebildet, da der Fokus mehr auf die Stabilität des Ratings gelegt wird.²¹

Die aus Kapitalmarktpreisen abgeleiteten Proxies für das Ausfallrisiko eines Schuldners sind hingegen

typischerweise „point-in-time“ (PIT)-Maße; dies bedeutet, dass sie auf Veränderungen des Konjunkturzustands reagieren und daher entsprechend stärker schwanken als TTC-Ratings.

Bei der Auswahl der Benchmark muss dieser Umstand bedacht werden. Handelt es sich bei dem zu validierenden Modell um ein TTC-Modell, bieten sich externe Ratings als Benchmark an. Bei einem PIT-Modell sollte hingegen ein aus Kapitalmarktpreisen abgeleitetes Proxy als Benchmark verwendet werden.²² Bei diesem ist jedoch zu beachten, dass aufgrund ihrer im Vergleich zu internen Ratings, die im Normalfall nur einmal pro Jahr aktualisiert werden, hohen Schwankung das Ergebnis der Untersuchung stark vom Stichtag der Untersuchung abhängen kann.

Neben dieser bis dato angesprochenen unterschiedlichen Ratingphilosophie sollten die verschiedenen Ratings denselben Risikoparameter abbilden. Es ist demnach etwa zu berücksichtigen, ob die Ratings ausschließlich als PD-Schätzer zu sehen sind oder ob sie auf den Expected

¹⁶ Ein CDS ist ein Vertrag zur Absicherung gegen Kreditrisiken, d. h., der Sicherungsgeber gibt dem Sicherungsnehmer die Zusage, Ausgleichszahlungen in Höhe eines eventuell eintretenden Verlusts zu übernehmen, wenn vorher festgelegte Kreditereignisse eintreten. Im Gegenzug zahlt der Sicherungsnehmer dem Sicherungsgeber eine Prämie über die Laufzeit der Absicherung, den sogenannten CDS-Spread (in Prozent des Nominalbetrags der Forderung), die um so höher ist, je höher die Wahrscheinlichkeit ist, dass das Kreditereignis eintritt.

¹⁷ Zhu (2004) hat gezeigt, dass diese beiden Maße sich langfristig weitest gehend identisch verhalten, kurzfristig jedoch voneinander abweichen können. Bei beiden Maßen ist zu beachten, dass ihre Höhe nicht nur durch das Ausfallrisiko, sondern auch durch weitere Faktoren wie Liquidität, Steuern oder von den Investoren verlangte Risikoprämien beeinflusst werden, siehe z. B. Elton et al. (2001) oder Amato und Remolona (2003).

¹⁸ Siehe Amato und Remolona (2003).

¹⁹ Siehe Cantor (2001) und Standard & Poor's (2006).

²⁰ Inwiefern die Ratings der großen Agenturen wirklich unabhängig vom Konjunkturzustand sind, wurde in verschiedenen empirischen Studien untersucht; siehe Nickell et al. (2000), Bangia et al. (2002), Amato und Furfine (2004) und Löffler (2006).

²¹ Siehe Fons et al. (2002).

²² Treacy und Carey (1998) haben mithilfe von Interviews entdeckt, dass es sich bei den internen Ratingmodellen von (großen amerikanischen) Banken typischerweise um PIT-Ratingmodelle handelt. Weber et al. (1999) haben zudem herausgefunden, dass die Ratings der Modelle größerer deutscher Kreditinstitute stärker schwanken als die externen Ratings der entsprechenden Schuldner, was ein Hinweis darauf sein könnte, dass die internen Modelle großer deutscher Banken eher PIT als TTC sind.

Loss abstellen. Zudem sollten sich die verschiedenen Ratings auf denselben Zeithorizont beziehen. Die Autoren sind sich der Tatsache bewusst, dass diese Anforderungen bei den herangezogenen Benchmarks nicht durchgängig erfüllt werden. Dennoch wird auf sie zurückgegriffen, da in der Praxis häufig keine „besseren“ Benchmarks für LDP zur Verfügung stehen.

Im folgenden Kapitel wird die Anwendung der dargestellten Methode anhand eines einfachen Beispiels demonstriert. Hierbei werden die Ergebnisse eines fiktiven bankinternen Ratingmodells für Staaten mit den Bonitätseinstufungen der drei großen Ratingagenturen Standard & Poor's (S&P), Moody's und Fitch sowie mit auf dem Kapitalmarkt beobachtbaren CDS-Spreads verglichen.

5 Beispiel zur Anwendung der Methode

In diesem Abschnitt wird die Anwendung der in Kapitel 4 vorgestellten Methode mittels eines Beispiels näher verdeutlicht. Hierzu werden die (fik-

tiven) Ergebnisse eines bankinternen Ratingmodells für Staaten zum einen den Ratings der Ratingagenturen S&P, Moody's und Fitch und zum anderen auf dem Markt beobachtbaren (CDS-)Spreads gegenübergestellt.

In Tabelle 1 werden die Ratings bzw. die CDS-Spreads der einzelnen Staaten dargestellt.²³ Bei den Ergebnissen des internen Ratingmodells handelt es sich um fiktive Werte auf einer Ratingskala von 1 bis 12 mit 1 als bestem Rating. Bei den herangezogenen CDS-Spreads handelt es sich um auf dem Kapitalmarkt beobachtete Werte.²⁴

Basierend auf diesen Ausgangsdaten wird in Anlehnung an die in Kapitel 4 beschriebene Vorgehensweise zur Berechnung von τ_x zunächst für jedes Merkmal (also für das interne Ratingsystem, die Ratings der drei Ratingagenturen sowie die CDS-Spreads) eine 9×9 -Matrix erstellt. In den Spalten und Zeilen der Matrix finden sich die jeweiligen Staaten (Kreditnehmer). In Tabelle 2 wird exemplarisch die Beurteilungsmatrix für das interne Ratingverfahren dar-

Tabelle 1

Ratingeinstufungen und CDS-Spreads der Staaten

Schuldner	Internes Ratingmodell	S&P	Moody's	Fitch	CDS-Spreads
Brasilien	8	BB+	Ba2	BB+	71
Ungarn	3	BBB+	A2	BBB+	19
Mexiko	4	BBB	Baa1	BBB	34
Polen	3	A-	A2	BBB+	8
Russland	5	BBB+	Baa2	BBB+	42
Südkorea	2	A-	A3	A+	16
Türkei	9	BB-	Ba3	BB-	148
Ukraine	9	BB-	B1	BB-	131
Venezuela	10	BB-	B2	BB-	251

Quelle: Standard & Poor's, Moody's, Fitch, Deutsche Bank (2007).

²³ Generell sollte zur Validierung ein möglichst großes Sample verwendet werden, damit die Ergebnisse nicht durch einzelne Ausreißer verzerrt werden. Um das Beispiel übersichtlich zu halten, werden im Folgenden jedoch nur die Ratings von neun Staaten betrachtet.

²⁴ Die Daten zu den CDS-Spreads (5-Jahres-CDS-Spreads) stammen von Deutsche Bank (2007); Stand: 6. Juli 2007.

Tabelle 2

Beurteilungsmatrix für das interne Ratingverfahren

	Brasilien	Ungarn	Mexiko	Polen	Russland	Südkorea	Türkei	Ukraine	Venezuela
Brasilien	0	-1	-1	-1	-1	-1	1	1	1
Ungarn	1	0	1	1	1	-1	1	1	1
Mexiko	1	-1	0	-1	1	-1	1	1	1
Polen	1	1	1	0	1	-1	1	1	1
Russland	1	-1	-1	-1	0	-1	1	1	1
Südkorea	1	1	1	1	1	0	1	1	1
Türkei	-1	-1	-1	-1	-1	-1	0	1	1
Ukraine	-1	-1	-1	-1	-1	-1	1	0	1
Venezuela	-1	-1	-1	-1	-1	-1	-1	-1	0

gestellt. Eine 1 in einer Zelle bedeutet, dass der Staat in der Zeile eine bessere oder die gleiche Ratingeinstufung zugewiesen bekommen hat wie der Staat in der Spalte. So weist etwa die Zeile Südkorea nur 1 auf, weil diesem Staat von dem internen Ratingverfahren die beste Einstufung der neun Staaten zugewiesen wurde. Eine -1 wird hingegen vergeben, wenn der Staat der Zeile ein schlechteres internes Rating aufweist als der Staat der Spalte. Auf der Diagonalen findet sich definitionsgemäß durchgehend eine 0.

Nachdem für jedes der fünf Merkmale die Beurteilungsmatrix erstellt wurde, wird die Matrix für das interne Ratingverfahren jeweils mit den anderen Matrizen multipliziert. In Tabelle 3 wird exemplarisch die Produktmatrix für das interne Ratingverfahren und die Matrix für Ratings

von S&P dargestellt. Eine 1 in einer Zelle ergibt sich immer dann, wenn die entsprechende Zelle in beiden Matrizen übereinstimmend eine 1 oder eine -1 aufweist. Dies bedeutet, dass beide Staaten, auf die sich die jeweilige Zelle bezieht, in beiden Merkmalen bezüglich ihrer Bonität nicht in entgegengesetzter Reihenfolge stehen. So erhält z. B. Südkorea sowohl vom internen Ratingverfahren als auch von S&P ein besseres Rating als Ungarn. Eine -1 in der Produktmatrix ergibt sich hingegen dann, wenn die Reihenfolge bei den beiden verglichenen Staaten bei den betrachteten Merkmalen entgegengesetzt ist. Demnach wird z. B. die Bonität von Russland vom internen Ratingverfahren schlechter eingestuft als jene von Mexiko, von S&P hingegen erhält Russland ein besseres Rating als Mexiko.

Tabelle 3

Produktmatrix für (fiktives) internes Rating und Rating von S&P

	Brasilien	Ungarn	Mexiko	Polen	Russland	Südkorea	Türkei	Ukraine	Venezuela
Brasilien	0	1	1	1	1	1	1	1	1
Ungarn	1	0	1	-1	1	1	1	1	1
Mexiko	1	1	0	1	-1	1	1	1	1
Polen	1	1	1	0	1	-1	1	1	1
Russland	1	-1	-1	1	0	1	1	1	1
Südkorea	1	1	1	1	1	0	1	1	1
Türkei	1	1	1	1	1	1	0	1	1
Ukraine	1	1	1	1	1	1	1	0	1
Venezuela	1	1	1	1	1	1	1	1	0

Quelle: Standard & Poor's.

Nachdem die entsprechenden Produktmatrizen gebildet wurden, kann für jede Produktmatrix die Kennzahl τ_x anhand Formel (1) berechnet werden. Für den Vergleich der Ratingeinstufungen des internen Rating-systems mit jenen von S&P berechnet sich τ_x zum Beispiel wie folgt:

$$\tau_x = \frac{58}{9(9-1)} \approx 0,81$$

In Tabelle 4 werden die (gerundeten) Ergebnisse für das vorliegende Beispiel festgehalten:

Es zeigt sich, dass im vorliegenden

Tabelle 4

(Gerundete) Ergebnisse für τ_x

Verfahren	τ_x
Internes Ratingverfahren und S&P	0,81
Internes Ratingverfahren und Moody's	0,86
Internes Ratingverfahren und Fitch	0,83
Internes Ratingverfahren und CDS-Spreads	0,89

fiktiven Beispiel das Ergebnis für alle vier Vergleiche Werte über 0,8 liefert, wobei das τ_x für das Risikomaß CDS-Spread am höchsten ist.

Bei der Beurteilung der Ergebnisse sind die oben angesprochenen Punkte, wie etwa die Ratingphilosophie, der abgebildete Risikoparameter oder der Zeithorizont zu berücksichtigen.

6 Fazit

In diesem Beitrag wurde eine mögliche Vorgehensweise dargestellt, mit der eine quantitative Validierung von Ratingmodellen für LDP durchgeführt werden kann. Eine notwendige Voraussetzung für ihre Anwendung ist das Vorhandensein eines geeigneten Vergleichsmaßes. Die in diesem Beitrag angeführten Vergleichsmaße „Externe Ratings“ und „Bond bzw. CDS-Spreads“ sind für typische LDP wie Staaten, Banken und große Unternehmen verfügbar, weswegen sich die dargestellte Methode besonders gut für diese LDP eignet.

Bei der Auswahl des Vergleichsmaßes ist zu bedenken, dass die Aussagekraft der Ergebnisse stark von der Qualität²⁵ des Vergleichsmaßes abhängt, da bei der dargestellten Methode nicht direkt die Qualität der Ergebnisse des internen Ratingmodells, sondern die Übereinstimmung von dessen Ergebnissen mit den Ergebnissen des Vergleichsmaßes beurteilt wird. Somit kann aus einem hohen Wert für τ_x nur dann auf eine hohe Trennschärfe des internen Ratingmodells geschlossen werden, wenn das Vergleichsmaß über eine hohe Trennschärfe verfügt. Umgekehrt bedeutet dies auch, dass aus einem niedrigen Ergebnis für τ_x nicht unmittelbar auf eine schwache Trennschärfe des internen Ratingmodells geschlossen werden sollte. Vielmehr sollten die Gründe für den niedrigen Wert, die beispielsweise in einer niedrigen Trennschärfe des Vergleichsmaßes liegen können, untersucht werden.

²⁵ Qualität im Sinne von Trennschärfe.

Literaturverzeichnis

- Amato, J. und C. Furfine. 2004.** Are Credit Ratings Pro-cyclical? In: Journal of Banking and Finance 28. 2641–2677.
- Amato, J. und E. Remolona. 2003.** The Credit Spread Puzzle. In: BIS Quarterly Review. Dezember. 51–63.
- Bangia, A., F. Diebold, A. Kronimus, C. Schagen und T. Schuermann. 2002.** Ratings Migration and the Business Cycle, with Application to Credit Portfolio Stress Testing. In: Journal of Banking and Finance 26. 445–474.
- BBA und ISDA – British Bankers’ Association, London Investment Banking Association und International Swaps and Derivatives Association. 2005.** Low Default Portfolios. Joint Industry Working Group Discussion Paper.
- BCBS – Basel Committee on Banking Supervision. 2005a.** Validation of Low-Default Portfolios in the Basel II Framework. Basel Committee Newsletter 6. September.
- BCBS – Basel Committee on Banking Supervision. 2005b.** Studies on the Validation of Internal Rating Systems. Working Paper 14.
- Cantor, R. 2001.** Moody’s Investors Service Response to the Consultative Paper Issued by the Basel Committee on Banking Supervision “A New Capital Adequacy Framework”. In: Journal of Banking and Finance 25. 171–185.
- CEBS – Committee of European Banking Supervisors. 2006.** Guidelines on the Implementation, Validation and Assessment of Advanced Measurement (AMA) and Internal Ratings Based (IRB) Approaches. April.
- Deutsche Bank. 2007.** Finanzmarktmonitoring – CDS Spreads. Deutsche Bank Research. <http://www.dbresearch.com/servlet/reweb2.ReWEB?rwkey=u727019> (Stand: 6. Juli 2007).
- Deutsche Bundesbank. 2003.** Validierungsansätze für interne Ratingsysteme. In: Monatsbericht September. Frankfurt am Main.
- Elton, E., M. Gruber, D. Agrawal und C. Mann. 2001.** Explaining the Rate Spread on Corporate Bonds. In: Journal of Finance 56. 247–277.
- Emond, E. und D. Mason. 2002.** A New Rank Correlation Coefficient with Application to the Consensus Ranking Problem. In: Journal of Multi-Criteria Decision Analysis 11. 17–28.
- Fons, J., F. Cantor und C. Mahoney. 2002.** Understanding Moody’s Corporate Bond Ratings and Rating Process. Special Comment. Moody’s Investors Services.
- Hornik, K., R. Jankowitsch, M. Lingo, S. Pichler und G. Winkler. 2006.** Validation of Credit Rating Systems Using Multi-Rater Information. Working Paper. November.
- Löffler, G. 2006.** Can Rating Agencies look Through the Cycle? Working Paper. Universität Ulm.
- Nickell, P., W. Perraudin und S. Varotto. 2000.** Stability of Rating Transitions. In: Journal of Banking and Finance 24. 203–227.
- OeNB und FMA – Oesterreichische Nationalbank und Finanzmarktaufsicht. 2004.** Leitfadenreihe zum Kreditrisiko: Ratingmodelle und -validierung. Wien.
- Standard & Poor’s. 2006.** Corporate Ratings Criteria. New York.
- Treacy, W. und M. Carey. 1998.** Credit Risk Rating at Large U.S. Banks. In: Federal Reserve Bulletin. November. 897–921.
- Weber, M., J. Krahen und F. Vossmann. 1999.** Risikomessung im Kreditgeschäft: Eine empirische Analyse bankinterner Ratingverfahren. In: Zeitschrift für betriebswirtschaftliche Forschung. Sonderheft 41. 117–142.
- Zhu, H. 2004.** An Empirical Comparison of Credit Spreads between the Bond Market and the Credit Default Swap Market. BIS Working Paper 160. Basel.