

Veröffentlicht in  
Kredit & Rating Praxis  
4/2004

“Ratingevidenz: Die Qualität von Ratingnoten”

Seite 22-24

Mit freundlicher Genehmigung der Kredit & Rating Praxis-Redaktion  
Akademischer Verlag St. Gallen, St. Gallen  
([www.krp.ch](http://www.krp.ch))



## Ratingqualität

# Ratingevidenz: Die Qualität von Ratingnoten

*Oliver Everling, Werner Gleißner*

Mit dem Basel II-Abkommen sowie die Mindestanforderungen an das Kreditgeschäft der Kreditinstitute haben sich Ratings auch in Deutschland zur «Visitenkarte» der Unternehmen entwickelt. Das Rating beschreibt aus Sicht eines Fremdkapitalgebers die wirtschaftliche Fähigkeit, rechtliche Bindung und Willigkeit eines Unternehmens, seinen zukünftigen Zahlungsverpflichtungen nachzukommen. Ein ausreichend gutes Rating ist für ein Unternehmen erforderlich, um den erforderlichen finanziellen Spielraum für Investitionen und Wachstum zu sichern – oder eine existenzbedrohende Kündigung des Kreditrahmens durch die Kreditinstitute zu verhindern (vergleiche Achleitner/Everling, *Handbuch Ratingpraxis*, 2004).

Die Entwicklung einer Ratingstrategie, also die gezielte Vorbereitung des Unternehmens auf das Rating, wird zunehmend zu einer der wesentlichen Aufgaben der Unternehmensleitung, um im Wettbewerb auf den Finanzmärkten zu bestehen (vergleiche Gleißner/Füser, *Leitfaden Rating*, 2. Auflage, 2003). Ausgangspunkt einer solchen Ratingstrategie ist das heutige Rating und die Prognose der zukünftigen Entwicklung des Ratings, die von der Unternehmensplanung und der mit ihr verbundenen Risiken abhängig ist (vergleiche Gleißner, *Rating-Prognosen am Beispiel einer Investitionsentscheidung*, *Rating Aktuell*, 6/2003) Eine solche Status-Quo-Beschreibung des Ratings kann eine Selbsteinschätzung des Unternehmens sein (zum Beispiel erstellt mittels Ratingsoftware), die bereits vorliegende Beurteilung durch ein Kreditinstitut oder eine Ratingagentur oder aber auch das «indikative Rating», das ein Rating Advisor zu Beginn eines Beratungsprojektes erstellt hat.

Da der Kreditrahmen (einer Bank), die Auftragserteilung (beim Rating eines Kunden) oder wesentliche Inhalte der zu erarbeitenden Ratingstrategie (bei einer Selbsteinschätzung des Ratings) von der zugrunde liegenden Ratingnote abhängt, ist dessen Qualität von grundlegender Bedeutung. Genau diese Qualität eines vorliegenden Ratings beschreibt die Ratingevidenz (englisch «rating evidence»). Die Evidenz gibt Auskunft über die Beweislage und Deutlichkeit des Ratings. Sie wird auf einer Skala von 100 Prozent (vollständige Evidenz) bis 0 Prozent (keine Evidenz) zum Ausdruck gebracht. In Anbetracht der qualitativ doch sehr unterschiedlichen Ratingverfahren (gleich, ob diese quantitative oder auch

qualitative Faktoren umfassen), dem unterschiedlichen Qualifikationsstand der involvierten Personen und der variierenden Verfügbarkeit von Daten unterscheidet sich offensichtlich die Qualität von Ratingnoten ganz erheblich. Mit der «Ratingevidenz» wird erstmalig ein Qualitätsstandard für Ratings erstellt – also quasi das Rating eines Ratings, ein «Meta-Rating».

Die Ratingevidenz reflektiert dabei die Beurteilung einer einzelnen vorliegenden Ratingnote. Es handelt sich nicht um eine grundsätzliche Beurteilung eines Ratingsystems oder einer Ratingagentur und auch nicht um einen Transparenzstandard, was einen Unterschied zu den DVFA-Ratingstandards darstellt<sup>1</sup>. Durch die Einschätzung der Ratingevidenz werden auch keine Zertifizierungen vorgenommen, wie sie durch die zu gründende Rating Cert GmbH zur Einhaltung von Mindestanforderungen im Rating geplant sind.

Die Evidenz von Ratings ein und derselben Ratingagentur kann nämlich in Abhängigkeit der jeweils eingesetzten Personen und der verfügbaren Daten erheblich schwanken. Die Ratingevidenz stellt dabei ein ex ante Qualitätsmaß dar, das heißt, die Beurteilung der Qualität des Ratings erfolgt ausschließlich anhand von Informationen, die zum Zeitpunkt des Ratings auch bereits zur Verfügung standen. Die Beurteilung der Ratingevidenz unterscheidet sich damit deutlich von einem so genannten «Backtesting» der Ratings einer Ratingagentur oder eines Kreditinstituts. Beim Backtesting wird zum einen nicht versucht, die Qualität einer einzelnen Ratingnote einzuschätzen. Stattdessen wird ein Qualitätsmaßstab für sämtliche Ratings einer Ratingagentur oder Bank (oder zumindest einer wesentlichen Teilmenge) beziehungsweise ihrer jeweiligen Ratingsysteme mittels statistischer Verfahren erstellt. Ein derartiges Backtesting basiert auf einem statistischen Vergleich prognostizierter Ausfallwahrscheinlichkeiten mit den später tatsächlichen eingetretenen. Neben statistischen Tests spielt für die Beurteilung von ganzen Ratingverfahren auch die Prüfung des Kreditgeschäfts der Kreditinstitute durch die Wirtschaftsprüfer (gemäß IDW/PS 522, Prüfung der Adressausfallrisiken und des Kreditgeschäfts von Kreditinstituten) eine Rolle. Auch dieser Prüfungsstandard, der sich zudem nur auf Kreditinstitute bezieht, hat einen anderen Fokus als die Prüfung der Ratingevidenz. Prüfungsgegenstand ist Wirksamkeit und Angemessenheit des internen Kontrollsystems (einschließlich der Ratingverfahren), also eine Systemprüfung und keine Einzelprüfung von



Ratings. Auch wenn selbstverständlich die Qualität des Ratingsystems eine wesentliche Determinante der Ratingevidenz ist, spielen bei einer derartigen Einzelbetrachtung eines Ratings – wie sie für ein geratetes Unternehmen maßgeblich ist – doch einige weitere, teils oben schon genannte, Faktoren eine Rolle<sup>2</sup>.

Grundsätzliches Problem jeder statistischen, auf historischen Daten basierenden Prüfung von Ratings besteht darin, dass offensichtlich zum Zeitpunkt des Ratings eine Qualitätseinschätzung so nicht möglich ist. Der Maßstab der Ratingevidenz schließt hier eine Lücke: Mit ihm ist schon zum Zeitpunkt der Erstellung eines Ratings eine Ersteinschätzung hinsichtlich der Ratingqualität möglich. Der Wert der Ratingevidenz bringt zum Ausdruck, inwieweit die Ratingnote als erwartungstreuer und konsistenter Schätzer der tatsächlichen Verteilungsfunktion der Ausfallwahrscheinlichkeit interpretiert werden kann. Vereinfacht ausgedrückt wird damit also abgeschätzt, in welcher Bandbreite um das angegebene Rating (beziehungsweise die zugehörige Ausfallwahrscheinlichkeit) die tatsächliche Ausfallwahrscheinlichkeit liegen wird.

Die Beurteilung der Ratingqualität bezieht sich dabei (im Allgemeinen) auf einen mittleren Prognose-Horizont (4- bis 5-Jahres-Zeitraum) und nicht – wie nach der Neuen Basler Eigenkapitalvereinbarung (Basel II) – auf eine 1-jährige Ausfallprognose. Gemessen wird die Ratingevidenz auf einer Skala von 0 bis 100: 0 Prozent bedeutet dabei ein rein zufälliges Ratingergebnis. Ein Evidenzwert von 100 Prozent steht für die beste (theoretisch mögliche) Beurteilung eines Ratings. Bei dieser Beurteilung würde bei einer (hypothetischen) unendlich großen Stichprobe von Ratingnoten dieser Güte im Mittel exakt die vorhergesagte Ausfallwahrscheinlichkeit eintreten.

Abhängig ist die Ratingevidenz von einer Vielzahl von Kriterien, die in einem heuristischen Bewertungsmodell miteinander verbunden sind. Die verwendeten Kriterien beziehen sich dabei auf

- > Organisation und Arbeitsabläufe bei der Erstellung des Ratings,
- > verwendete Ratingmethode,
- > eingesetztes Ratingsystem und Verfahren,
- > verfügbare und genutzte Daten,
- > erreichte Zertifizierung von Ratinggesellschaften und Ratingsystemen,
- > Reputation sowie
- > eingesetzte Ressourcen (insbesondere Qualifikation und Erfahrung der eingesetzten Personen).

Die hier aufgeführten Kriterien zeigen die Notwendigkeit einer sehr differenzierten Beurteilung einzelner Ratings. So ist beispielsweise selbst bei einem hervorragend geeigneten Ratingverfahren eine gute Ratingevidenz nicht sichergestellt, wenn der beurteilenden Ratingagentur (oder dem Rating Advisor) ausschließlich öffentlich zugängliche Informationen zugänglich waren, wenn diese systembedingt nicht ausreichen. Auch eine wissenschaftlich bestens anerkannte Ratingmethodik führt zu einer deutlich schwächeren Einschätzung der Evidenz eines Ratings, wenn die erforderlichen Arbeitsprozesse wegen mangelhafter Dokumentation nicht stringent eingehalten werden oder die durchführenden Mitarbei-

ter über keine adäquate Qualifikation oder Erfahrung verfügen. Umgekehrt ist es selbst für erfahrene Mitarbeiter einer Ratingagentur grundsätzlich unmöglich, Ratings mit einer hohen Evidenz zu erzeugen, wenn die von ihnen angewandte Verfahrensweise methodische Fehler aufweist oder bei statistischen Analysen die Grundgesamtheit der ausgewerteten Daten zu gering oder nicht repräsentativ ist.

Mit dem Konzept der Ratingevidenz wurde dem Finanzausschuss des Deutschen Bundestages zu seiner Sitzung am 3. März 2004 eine Alternative zur pauschalen Anerkennung oder Nicht-Anerkennung von Ratingagenturen vorgestellt. Auch Analysten sonst «anerkannter» Ratingagenturen können sich in Einzelfällen irren. Für viele Unternehmen fehlen sogar den führenden Agenturen geeignete analytische Ressourcen, um zu treffsicheren Beurteilungen zu gelangen. So gibt es durchaus Fälle, in denen die Ratingagenturen die Annahme eines Ratingauftrages ablehnen, da sie sich nicht im Besitz der dafür notwendigen analytischen Ressourcen sehen. Nehmen sie dennoch den Ratingauftrag an, kann das Ergebnis nur von geringer Evidenz sein.

Auch wurden einzelne Fälle bekannt, in denen die mit bestimmten Ratings betrauten Analysten aus unterschiedlichen Gründen abgezogen wurden, ohne dass bei den Agenturen bereits adäquater Ersatz für die entfallenden Kapazitäten gefunden worden wäre. Angesichts der wichtigen Rolle einzelner Analysten in den Verfahrensweisen der führenden Ratingagenturen muss ein solcher Abzug entscheidender Ressourcen zu einer Verringerung der Ratingevidenz führen. Wenn Emittenten nicht einmal mehr einen Ansprechpartner bei «ihrer» Ratingagentur finden, kann dies ein Indiz für eine geringe Evidenz des Ratings sein. Wo keine Ressourcen zur laufenden Beobachtung von Unternehmen und zur Evaluation von Einflüssen auf Bonitätsveränderungen zur Verfügung stehen, können auf Dauer keine Ratings mit hoher Evidenz ermittelt werden.

Der Begriff der Ratingevidenz korrespondiert in einer ex post Betrachtung von Ratings mit dem Korrelationskoeffizienten, dem Gini-Koeffizienten der Lorenzkurve und anderen Korrelations- und Konzentrationsmaßen. So ist beispielsweise der Korrelationskoeffizient normiert auf -1 bis +1. Eine hohe Korrelation der tatsächlich eintretenden Ausfallraten mit den durch Ratings prognostizierten Ausfallwahrscheinlichkeiten deutet auf eine hohe, historisch begründete Evidenz erteilter Ratings hin. Besteht zwischen (historischen) Ausfallhäufigkeiten und (prognostizierten) Ausfallwahrscheinlichkeit eine Korrelation von Null, dürfte dies in der Regel – je nach Größe der Stichprobe beziehungsweise der Grundgesamtheit – ein Indiz für eine sehr niedrige Evidenz sein, im Extremfall einer Evidenz von Null. Dass Ratingsysteme meist nur für einen mehr oder weniger großen Anteil aller Fälle zu korrekten Ergebnissen gelangen, ist in den Wissenschaften unstrittig. So sind für die Systeme des Kreditratings in den Wirtschaftswissenschaften zufällige und auch systematische Fehler bekannt. So wird der Fehler eines «zu guten» Ratings als Alpha-Fehler (in anderen Quellen als Fehler 1. Art oder 1. Ordnung bezeichnet),



der systematische Fehler eines «zu schlechten» Ratings als Beta-Fehler identifiziert (in anderen Quellen als Fehler 2. Art oder 2. Ordnung bezeichnet). Während aus dem mathematisch-statistischen Modell selbst im Einzelfall der Fehler nicht zu erkennen sein mag, so kann durch eine ganzheitliche Betrachtung durchaus bestimmt werden, ob ein erteiltes Rating als «korrekt» – im Sinne einer hohen Evidenz – beurteilt werden kann.

Dies sei an einem konstruierten Beispiel verdeutlicht: Würde der Papst ein Darlehen in der deutschen Filiale eines Kreditinstituts nachfragen, würde er aller Wahrscheinlichkeit nach durch das bankinterne Ratingsystem abgelehnt, wenn er nicht als Papst erkannt wird. Als 84-jähriger, lediger, polnischer Emigrant ohne Nachfolgeregelung, mit sonst unbekanntem Arbeitsverhältnis beziehungsweise Geschäftsmodell und der Absicht, den Kreditbetrag aus der Bundesrepublik Deutschland in den Vatikanstaat zu verschaffen, würde Johannes Paul II bei deutschen Instituten auf Basis der Modellannahmen wohl kaum Kredit erhalten. Sein Alter, sein Familienstand, die unsichere Antwort auf die Frage nach seinem Nachfolger und so weiter sprächen nach dem Kriterienkatalog des Computermodells zunächst einmal gegen ein gutes Rating, ganz zu schweigen von dem durch das Programm abgefragten Warnsignal «körperlicher Gebrechen». Hierbei handelt es sich um einen Fehler 2. Art, da der Papst wohl unstrittig guter Bonität ist. Die Evidenz des bankintern erteilten Ratings ist in diesem Fall gering, obwohl für die Mehrzahl der typischen Kunden der Bankfiliale das Rating eine hohe Evidenz aufweisen dürfte. In diesem speziellen Fall hat ein an sich sinnvolles Ratingsystem hier relevante Informationen nicht ausgewertet.

Es erübrigt sich hier, neben dem Fall des Fehlers 2. Art analog auf den Fehler 1. Art hinzuweisen. Zahlreiche Insolvenzen zeugen davon, wie zu gut eingestufte Kreditnehmer später zu Abschreibungen führten. Milliardenkredite an mittellose Immobilienspekulanten wie auch Kleinkredite an Gewerbetreibende zeugen davon, dass neben korrekten Klassifizierungen stets auch Fehleinschätzungen zu beobachten sind. Da aber meist nur ex post, kaum aber ex ante sicher von einer Fehlklassifizierung gesprochen werden kann, liefert der Begriff der Ratingevidenz ein wichtiges Element für den notwendigen terminologischen Rahmen.

Mit dem Gütemaßstab «Ratingevidenz» besteht für jeden an Ratings Interessierten die Möglichkeit, eine vorliegende Ratingnote selbst einzuschätzen. Um diese Einschätzung von Ratings möglichst effizient zu gestalten, werden beispielhafte Ratingnoten, für die die Evidenz bereits bestimmt wurde, im Internet verfügbar gemacht (vergleiche [www.rating-evidence.de](http://www.rating-evidence.de)). Zusätzlich besteht für interessierte Anwender die Möglichkeit, über die genannte Internet-Seite die kostenlose Software «Rating-Evidence-light» zu beziehen, die von der Rating Evidence GmbH gemeinsam mit der FutureValue Group AG entwickelt wurde. Diese Software hilft durch vorgegebene Orientierungsfragen, die Evidenz eines vorliegenden Ratings, sei es von einer Ratingagentur, einem Kreditinstitut oder einem Rating Advisor, einzuschätzen. Vor allem ist es jedoch möglich, vorab zu erkennen, ob die

vorliegende Ratingeinschätzung überhaupt eine geeignete Grundlage beispielsweise für die Entwicklung einer Ratingstrategie darstellt. Auch für die kritische Diskussion der Ratingeinschätzung durch ein Kreditinstitut, wie sie regelmäßig Thema der sogenannten «Bilanzgespräche» ist, bietet ein derartiges Hilfsmittel vielfältige Anknüpfungspunkte.

Insgesamt soll durch die Beurteilung der Evidenz von Ratings ein Beitrag geleistet werden, um die Qualitätsstandards der Ratings zu verbessern. Unternehmerische Fehlentscheidungen auf Grundlage falscher Ratings sollen so in ihrer Wahrscheinlichkeit reduziert werden. Durch die Verbesserung der Qualität der Ratings in Deutschland kann zudem erreicht werden, die Akzeptanz der Ratings zu verbessern und sie zu einem der wichtigsten Instrumente für die Kapitalmarktkommunikation zu etablieren. <

## Fussnoten

- 1 Gemäß dem Basel II-Akkord wird dabei unter «Rating-System» der Überbegriff zu Methoden, Prozessen, Kontrollen, Datenerhebung und IT-Instrumenten verstanden, die zur Bestimmung von Kreditrisiken, der Zuordnung der Ratings und der Quantifizierung von Ausfall- und Verlustschätzungen dienen.
- 2 Neben der Prozess- beziehungsweise der Verfahrensprüfung spielt sowohl für die Ratings der Kreditinstitute als auch diejenige der Rating-Agenturen eine ergebnisorientierte Prüfung durch ein so genanntes «Backtesting» eine besondere Rolle, bei der insbesondere die Trennschärfe der eingesetzten Verfahren und die Erwartungstreue der Schätzer zur Ausfallwahrscheinlichkeit (und damit des Ratings) untersucht werden. Eine derartige Beurteilung umfasst statistische Kennzahlen wie die Alpha- und Beta-Fehler, den Gini-Koeffizient, die Entropie oder den Brier-Score.



## Autoren



Dr. Oliver Everling, Geschäftsführer, Rating Evidence GmbH und Geschäftsinhaber der Everling Advisory Services, Frankfurt am Main



Dr. Werner Gleißner, Vorstand FutureValue Group AG und Geschäftsführer RMCE RiskCon GmbH & Co. KG, Leinfelden-Echterdingen.  
E-Mail: [w.gleissner@futurevalue.de](mailto:w.gleissner@futurevalue.de)

Auszug aus Kredit & Rating Praxis 4/2004  
© Copyright by Akademischer Verlag St.Gallen AG, 2004