

# Compass-Benchmark

## White Paper

# Inhalt

1	Zusammenfassung .....	3
2	Datenbestand .....	4
3	Modellentwicklung .....	7
3.1	Generelles Vorgehen .....	7
3.2	Datenaufbereitung .....	8
3.3	Einzelfaktoranalyse .....	9
3.4	Multifaktoranalyse .....	12
3.5	Kalibrierung auf eine mittlere Insolvenzwahrscheinlichkeit .....	14
4	Validierung .....	17
4.1	Trennschärfe .....	17
4.2	Stabilität .....	18
4.3	Kalibrierung .....	19
5	Schlussbemerkung .....	21
6	Anhang .....	22
6.1	Literatur .....	22
6.2	Abbildungsverzeichnis .....	22
6.3	Tabellenverzeichnis .....	22

# 1 Zusammenfassung

Beim Compass-Benchmark handelt es sich um ein Ratingmodell österreichischer Firmen, das basierend auf Informationen des Jahresabschlusses aus dem Firmenbuch die Einjahres-Insolvenzwahrscheinlichkeit berechnet. Die durch das Modell errechnete Ratingnote entspricht also der Wahrscheinlichkeit, dass ein Unternehmen innerhalb eines Zeithorizonts von einem Jahr seinen Zahlungsverpflichtungen nicht nachkommen kann.

Bei der Entwicklung des Ratingmodells kamen modernste statistische Methoden zum Einsatz, die, ebenso wie die Datenbasis, den strengen Anforderungen an ein Risikomessverfahren genügen. Die Methodik ist dieselbe, welche bei den führenden europäischen Banken in der Beurteilung der Bonität ihrer Kunden eingesetzt wird, bei deutlich breiterer (weil gesamt-österreichischer) Datenbasis. Dadurch kann eine hohe Ähnlichkeit der Ergebnisse des Compass-Benchmark mit den Ratingverfahren führender Banken sichergestellt werden.

Der Compass-Verlag sieht es als wesentliches Qualitätsmerkmal eines Ratingmodells an, dass die methodischen Grundlagen der Modellentwicklung, die für die Entwicklung des Modells verwendeten Daten, sowie die Ergebnisse der empirischen Tests zur Messung der Leistungsfähigkeit so weit wie möglich offen gelegt und transparent gemacht werden. Dieses Dokument enthält daher Informationen

- über den Datenbestand, der für die Entwicklung und Validierung des Ratingmodells verwendet wurde
- über die angewandte Methodik, inklusive Kalibrierung auf eine marktgerechte Ausfallwahrscheinlichkeit
- über die Interpretation der im Modell eingesetzten Kennzahlen hinsichtlich Aussage und Risikogehalt
- über die Leistungsfähigkeit des Ratingsystems (Trennschärfe und Stabilität)

Dieses Dokument ermöglicht es, die Entwicklung und die Validierung des Modells konzeptionell nachzuvollziehen, und anhand der statistischen Angaben die Güte des Modells zu beurteilen. Zusätzlich zu diesem Dokument beantworten wir gerne Ihre Fachfragen.

## 2 Datenbestand

Der Datenbestand, der für die Entwicklung und Validierung des Compass-Benchmark verwendet wurde, stammt aus einer mehrjährigen Historie von Firmenbuch-Daten. Zum Monatsultimo Oktober 2013 wurden alle im Firmenbuch elektronisch eingelesenen Bilanzen bezogen und mit den Stammdaten (OENACE-Code, Rechtsform, Anzahl der Mitarbeiter, Postleitzahl des Unternehmenssitzes, ...) der bilanzierenden Firmen angereichert. Zusätzlich wurde für alle Unternehmen erhoben, ob und ggfs. wann für diese eine Insolvenz beantragt wurde. Aufgrund der gesetzlichen Einreichpflicht ans Firmenbuch repräsentiert der Entwicklungsdatenbestand daher einen maximal breiten Durchschnitt aller österreichischer Unternehmen, die Firmenbuch-bilanzpflichtig sind.

Insgesamt konnten knapp 970.000 Bilanzen zu knapp 160.000 Unternehmen bezogen werden. Im Schnitt findet sich also jedes Unternehmen 6,1-mal mit einem Jahresabschluss wieder, was es prinzipiell möglich macht, für ein einzelnes Unternehmen auch Vergleiche über verschiedene Jahre hinweg anzustellen. Die Stichtage der Bilanzen stammen überwiegend aus dem Zeitraum 2004-2012, wie nachfolgende Verteilung zeigt.

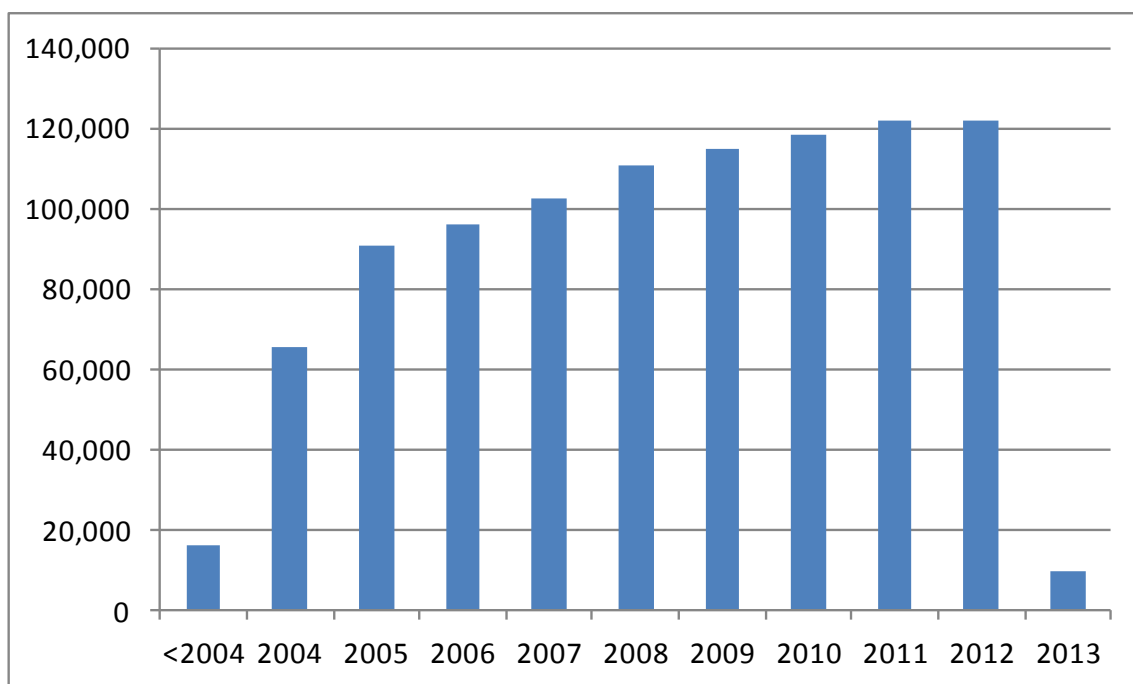


Abbildung 1: Zeitliche Staffelung der Bilanzen

Ein Insolvenzkennzeichen wiesen 4,6 % der Unternehmen auf. Die insolventen Unternehmen können für die Modellentwicklung vor ihrem Insolvenzdatum mehrfach Bilanzen eingereicht haben. Die Bilanzen nach dem Insolvenzstichtag sind allerdings nicht mehr relevant für die Modellentwicklung, weil diese Bilanzen keinen Prognosecharakter mehr besitzen.

Die Breite der Datenfelder orientiert sich am Eingabeschema der Firmenbuch-Bilanz. Theoretisch ist die Liste an lieferbaren Bilanzpositionen sehr lang, in der Praxis orientiert sich jedoch der Großteil der Unternehmen am Minimalschema, hier werden nur rund ein Dutzend Bilanzpositionen abgelegt. Nach Durchsicht der Belegungsquoten der einzelnen Bilanzpositionen wird klar, dass für ein umfassendes Ratingmodell nur eine relativ geringe Anzahl an Positionen wirklich verwendbar ist – der große Rest muss ungenutzt bleiben, weil

diese Positionen nur von sehr wenigen Firmen geliefert werden. Nachstehende Abbildung zeigt das "Rumpfschema" nach HGB.

AKTIVA	HGB_224_2	PASSIVA	HGB_224_3
Anlagevermögen	HGB_224_2_A	Eigenkapital (Negatives Eigenkapital)	HGB_224_3_A
Immaterielle Vermögensgegenstände	HGB_224_2_A_I	Kapitalrücklagen	HGB_224_3_A_II
Sachanlagen	HGB_224_2_A_II	Gewinnrücklagen	HGB_224_3_A_III
Finanzanlagen	HGB_224_2_A_III	Bilanzgewinn (Bilanzverlust)	HGB_224_3_A_IV
Umlaufvermögen	HGB_224_2_B	davon Gewinnvortrag/Verlustvortrag	HGB_224_3_A_IV_x
Vorräte	HGB_224_2_B_I	Unversteuerte Rücklagen	HGB_224_3_B
Forderungen und sonstige Vermögensgegenstände	HGB_224_2_B_II	Rückstellungen	HGB_224_3_C
Wertpapiere und Anteile	HGB_224_2_B_III	Verbindlichkeiten	HGB_224_3_D
Kassenbestand, Schecks, Guthaben bei Kreditinstituten	HGB_224_2_B_IV	Rechnungsabgrenzungsposten	HGB_224_2_C
Rechnungsabgrenzungsposten	HGB_224_2_C		

Abbildung 2: Bilanzschema mit HGB-Kennungen

Bezüglich des Berichtsschemas ist insbesondere festzuhalten, dass der Großteil der Bilanzen ohne Gewinn- und Verlustrechnung (GuV) vorliegt. Diese Positionen sind für die meisten Unternehmen nicht vorrätig, sodass die Ratingmodellierung darauf verzichten muss. Als elegante Annäherung an die Flußgrößen einer GuV hat sich hier allerdings die Möglichkeit erwiesen, Vergleiche von Bilanzpositionen zwischen den einzelnen Jahresabschlüssen zu ziehen. So kann etwa über die Veränderung des Eigenkapitals eine Kennzahl konstruiert werden, die hohe Korrelation zum Jahresüberschuss aufweist – und dergestalt können auch Bilanzkennzahlen approximiert werden, die normalerweise auf GuV-Positionen beruhen (etwa: Eigenkapitalrendite).

Entsprechend den Anforderungen an die Entwicklung von Ratingmodellen (s. Leitfaden der OeNB) wurde der Gesamtdatensatz in ein Entwicklungs- und Validierungssample unterteilt. Das Validierungssample, das eine Zufallsstichprobe von einem Drittel aller Bilanzen umfasst, wird für den statistischen Test des Modells verwendet (Trennschärferechnung), während die zwei Drittel des Entwicklungssamples für das Bauen des Modells zur Verfügung stehen. Nachstehende Abbildung gibt einen Überblick über das Mengengerüst.

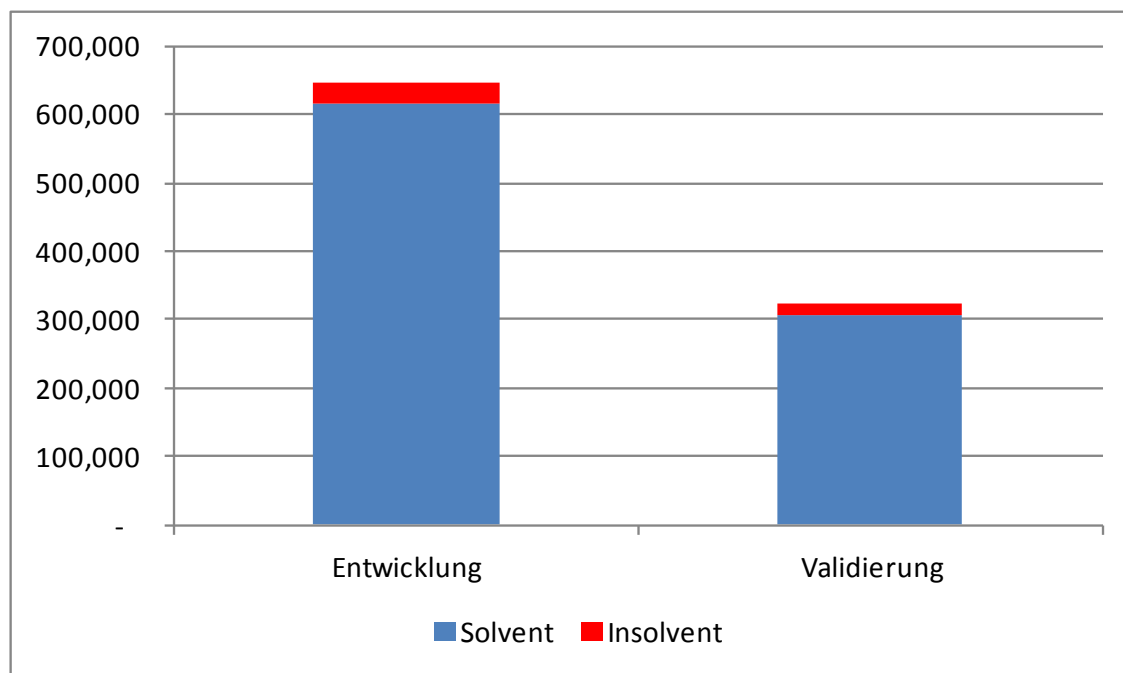


Abbildung 3: Mengengerüst des Gesamtdatenbestandes

Insgesamt ist die vorliegende Datenbasis im Vergleich zu führenden Banken im deutschsprachigen Raum eine exzellente Grundlage für die statistische Entwicklung eines Ratingmodells. Durch den hohen Abdeckungsgrad des Firmenbuches mit einem standardisierten Bilanzschema, und durch die vom Compass-Verlag automatisierte, transparente und konsistente Verarbeitung der Bilanzen, weist die Datenbasis zum Compass-Benchmark ein absolut hohes Qualitätsniveau auf. Weil auch in Hinkunft das Bilanzierungsschema in der modellierten Form Bestand haben wird, und die Bilanzen mittlerweile zum Großteil elektronisch eingereicht werden, entspricht die Anwendung des Ratingmodells auch vollkommen der Modellierung – ein Qualitätsvorteil im Datenbestand, den nur wenige Ratingmodelle für sich in Anspruch nehmen können.

## 3 Modellentwicklung

### 3.1 Generelles Vorgehen

Die Entwicklung des Compass-Benchmark orientiert sich an den Standards führender Banken und den Methoden des modernen Kreditrisikomanagements. In insgesamt 6 Stufen wird, ausgehend von einer Datenbasis historischer Bilanzen, eine Formel entwickelt, die für eine jede Bilanz ein Rating errechnen lässt. Die erste Stufe, Datenerhebung, gilt dabei als vorbereitender Schritt, und die letzte Stufe, Dokumentation, als begleitende Aktivität. Die mittleren 4 Stufen stellen die Modellentwicklung im engeren Sinne dar.

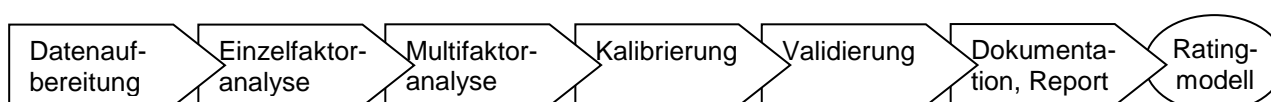


Abbildung 4: Entwicklungsprozess Compass-Benchmark

**Datenaufbereitung:** Bereits bei der Datenerhebung wird Wert darauf gelegt, alle Informationen auf Plausibilität und Qualität zu prüfen. Nichtsdestotrotz laufen die Daten ins Firmenbuch i.d.R. ohne manuelle Nachbearbeitung ein, sodass sich im Rohdatenbestand immer noch Fehler und Lücken in den einzelnen Variablen finden. Daher werden nach Einlesen der Rohdaten in einer Reihe von Aufbereitungsschritten Bereinigungen vorgenommen, welche die Datengrundlage in ihrer Qualität optimieren und für eine statistische Analyse vorbereiten. Zu diesen Bereinigungsverfahren gehören die Prüfung der Anforderungen an Mindeststandards, welche ein Datensatz einhalten muss, um in die Modellentwicklung mit aufgenommen werden zu können, eine gründliche Ausreißerdetektion, sowie die Imputation von fehlenden Werten. Endprodukt dieses Schrittes ist ein Datenbestand, der die höchsten Qualitätskriterien erfüllt, die mit rein statistischen (also nicht manuellen) Aufbereitungsmethoden zu erreichen sind.

**Einzelfaktoranalyse und -transformation:** Nach erfolgter Datenaufbereitung werden die einzelnen Kennzahlen (Bilanzkennzahlen sowie Stamminformationen wie Branche oder Standort) definiert und ihm Rahmen der sogenannten Einzelfaktoranalyse einer genauen Betrachtung unterzogen. Hierbei werden potenziell risikorelevante Faktoren gebildet, die für das Modell insoweit in Frage kommen, als dass sie auf neuen Bilanzen ebenso verfügbar sein müssen und fachliche Plausibilität sowohl als auch inhaltliche Akzeptanz aufweisen müssen. Für jeden Faktor erfolgt daraufhin die Untersuchung, wie seine Ausprägungen mit der Insolvenzwahrscheinlichkeit korrelieren. Dabei verstärkt eine mathematische Transformation der Faktoren den Informationsgehalt und bietet durch eine Glättung der Risikoverläufe eine optimale Grundlage für die sich anschließende Multifaktoranalyse. Endprodukt dieses Schrittes ist eine Menge an Risikofaktoren, deren einzelne Wirkungen auf die Insolvenzwahrscheinlichkeit ausgewiesen werden können.

**Multifaktoranalyse:** Die in der Einzelfaktoranalyse definierten Kennzahlen werden im Rahmen der Multifaktoranalyse mit Hilfe multivariater Regressionstechniken kombiniert. Dabei werden die in Kombination trennschärfsten Variablen identifiziert und ein Algorithmus entwickelt, der die Insolvenzwahrscheinlichkeit, basierend auf diesen Faktoren, berechnet. Die Kombination der Variablen unterliegt dabei auch einer fachlich-inhaltlichen Kontrolle, um die Ausgewogenheit der finalen Ratingformel zu garantieren. Endprodukt dieses Schrittes ist eine Menge an Risikofaktoren, die im Zusammenspiel eine möglichst trennscharfe Rangfolge an Bilanzratings errechnen lassen.

**Kalibrierung:** Eine über die Multifaktoranalyse hinausgehende Kalibrierung des Ratingmodells ist notwendig, da die mittlere Insolvenzwahrscheinlichkeit im Entwicklungsdatensatz nicht zwangsläufig mit der realen, am Markt beobachteten mittleren Insolvenzwahrscheinlichkeit übereinstimmt. Diese Abweichung wird im Rahmen der Kalibrierung korrigiert und die Insolvenzwahrscheinlichkeit auf ein in Österreich zu erwartendes Niveau bringt. Mit diesem Schritt steht die eigentliche Ratingformel, die nun für beliebige neue Bilanzen eine Insolvenzwahrscheinlichkeit auswerfen kann, welche im Mittel der Insolvenzwahrscheinlichkeit der österreichischen Unternehmen entspricht.

**Validierung:** Im diesem Schritt wird die Vorhersagekraft des Modells auf Basis eines separaten, von der Entwicklungsstichprobe unabhängigen Datensatzes validiert. Hierbei werden die Stabilität, die Trennschärfe (= Rangordnungsfähigkeit) und die Kalibrierung des Modells getestet. Die Validierung ist die Nagelprobe des Modells, weil das Modell hier zum ersten Mal auf neue, vorher noch nicht verwendete Bilanzen angewendet wird und dort seine Prognosefähigkeit beweisen muss. Endprodukt dieses Schrittes ist ein Maß für die Güte des Modells, sowie das Prüfsiegel, das die vertestete Modellformel in Einsatz gehen kann.

**Dokumentation:** Um für den Anwender des Ratingmodells, den mit den Ratingergebnissen konfrontierten Dritten, sowie den betroffenen Unternehmen selber die Entstehung und Qualität des Modells nachvollziehbar zu machen, wird die Entwicklung des Ratingverfahrens gründlich dokumentiert. Hierzu dienen dieses Dokument sowie die Programm-Codes der statistischen Analysen. Darüber hinaus gibt der Benchmark-Report detaillierte Einsicht in das Ratingmodell, indem anhand einer ausgewählten Bilanz alle Kennzahlen separat ausgewiesen werden und mehrjährig sowie im Branchen-Benchmark dargestellt werden.

Im nachfolgenden werden die Schritte Datenaufbereitung, Einzelfaktoranalyse, Multifaktoranalyse und Kalibrierung näher erläutert.

## 3.2 Datenaufbereitung

Die Verfahren zur Bestimmung von Insolvenzwahrscheinlichkeiten sind in der Fachliteratur gut untersucht und gelten in der Theorie als methodisch unkritisch. In der Praxis treten jedoch die typischen Schwierigkeiten im Umgang mit realen Daten auf: Fehlende Werte, Ausreißer, Autokorrelation und Multikollinearität, sowie nicht-normale Verteilungen. Um stabile Schätzergebnisse zu erhalten, ist es daher unabdingbar, nach Berechnung der Kennzahlen auf Rohdatenbasis einige Bereinigungen vorzunehmen, und zwar sowohl bei der Entwicklung des Ratings, als auch bei dessen Einsatz unter Anwendungsbedingungen.

**Fehlende Werte:** Datenlücken werden zunächst daraufhin untersucht, ob sie zufällig zustande kamen oder systematisch fehlen. Außerdem ist es für die Aufbereitung maßgeblich, ob die Anwendung des Ratingmodells fehlende Werte zulassen oder ausschließen sollte (sog. Mussfeld-Bedingung). In ersterem Fall ist es möglich, den fehlenden Wert als zulässige Ausprägung zu definieren, in letzterem Fall muss eine Befüllung der Fehlstelle mit einem validen Wert erfolgen (sog. Imputation), um den Datensatz nutzbar zu machen. Prinzipiell ist diese Imputation aber nur sinnvoll, wenn die Fehlstellen nicht überhand nehmen. Faktoren, die nur in einer kleinen Minderheit der Fälle gefüllt sind, werden daher in der Modellierung besser außen vor gelassen. Im Falle des Compass-Benchmark betrifft das den ganzen Bereich der GuV-Positionen, weil diese i. d. R. nur für große Unternehmen vorgelegt werden müssen, in der Rumpfbilanz der kleinen und mittleren Unternehmen aber nicht eingefordert werden. Diese schwach befüllten (<10 % der Fälle) Faktoren werden im Schritt Datenaufbereitung ausgesiebt.

**Ausreißerdetektion:** Durch die Definition von Kennzahlen als Quotienten, deren Zähler und Nenner aus weit variierenden Wertebereichen stammen können, finden sich in den meisten Einzelfaktoren große Ausreißer – also Werte, die den Kernbereich der fachlich erwarteten Zahlen deutlich über- oder unterschreiten. Dabei muss es sich nicht zwangsläufig um einen Fehler handeln. Bilanzen gehorchen nicht immer einer starren Stabilität, und wenn ein Unternehmen von einem Jahr auf das nächste größere Umstrukturierungen tätigt, kann sich



das in zunächst unrealistisch anmutenden Werten äußern. Um auch diese Fälle für den Datensatz nutzbar zu machen, werden Wertebereiche üblicherweise auf fachliche Glaubwürdigkeit hin begrenzt (Untergrenze=Floor, Obergrenze=Cap) und die unter- bzw. überschreitenden Werte auf die Floors und Caps hinauf bzw. herunter gesetzt.

**Autokorrelation und Multikollinearität:** Werden in das Regressionsverfahren zu viele sich ähnelnde Beobachtungen (also etwa: immer wieder die gleiche Bilanz, ohne Veränderung) eingespielt, fehlt es am Ende an Varianz, die durch das Regressionsverfahren erklärt werden sollte. Es kommt zu unrealistischen Fehlerabschätzungen und durch einzelne Beobachtungen leicht zu verzerrenden Schätzergebnissen. Dieses unter Autokorrelation bekannte Problem kann auch im betrachteten Fall auftreten, weil Bilanzen über die Zeit mehrfach eingespielt werden. Mit einem Hypothesentest (Durbin-Watson-Test) wurde jedoch überprüft, dass die Verzerrung durch Autokorrelation nicht abträglich ist. Ähnlich verhält es sich mit dem Problem, das auftritt, wenn der Regression zu viele gleichartige Variablen (also Einzelfaktoren) angeboten werden. Auch dafür wurden Tests zur Detektion entwickelt, die im Falle des Compass-Benchmark keine signifikanten Probleme der sog. Multikollinearität bezeugen.

**Normalverteilungsannahme:** Das angewandte Verfahren der logistischen Regression ist üblicherweise sehr robust gegenüber Verletzungen der Normalverteilungsannahme. Nichtsdestotrotz wurden im Rahmen der Datenaufbereitung Transformationen von einzelnen Variablen vorgenommen, vor allem deshalb, um Nichtlinearitäten besser abbilden zu können (s. dazu mehr im nächsten Abschnitt).

Im Prinzip wirken die Aufbereitungsschritte, die die Rohkennzahlen durchlaufen, wie ein intelligenter Filter – ein solcher füllt Datenlücken, glättet Verläufe, indem er Ausreißer ausschließt, skaliert die Kennzahl auf brauchbare Wertebereiche, und steigert somit den Informationsgehalt der Rohkennzahl für die Nutzbarkeit in einem Regressionsverfahren. Unter den Begriff Datenaufbereitung fallen aber auch viele einfachere Aufgaben wie das Extrahieren der Branchen aus dem OENACE-Code, das Zuspielden der Bundesländer zu den Postleitzahlen, das Detektieren des Insolvenzzeichens, das Errechnen von zeitlichen Veränderungsdaten, oder das Ziehen der Entwicklungs- bzw. Validierungsstichprobe.

### 3.3 Einzelfaktoranalyse

Der erste Schritt der Einzelfaktoranalyse ist die Definition von potenziellen Risikofaktoren. Dies sind zum einen Bilanzanalysekennzahlen, die sich in der Literatur finden bzw. selbsttätig konstruieren lassen, sowie zum anderen Stamminformationen wie die Branche, die Rechtsform oder die Größe des Unternehmens. Am Beispiel der letzteren Variable sei die Definitionsarbeit erläutert: Die Größe eines Unternehmens lässt sich z. B. mittels Umsatz messen, oder es wird die Bilanzsumme als Maß herangezogen, oder etwa auch der Mitarbeiterstand. All diese Größen können dann zu einem Stichtag bezogen werden, oder aber wird ein Durchschnitt der Größen über die letzten Jahre errechnet, oder die Größe wird in Bezug zu einer Benchmark gesetzt, oder in Klassen kategorisiert, oder durch eine Funktion (wie etwa dem Logarithmus) transformiert. All dies lässt sehr viele Freiheitsgrade offen, die Methodik der Ratingmodellierung gibt hier kaum Grenzen vor.

Um eine Stufe komplexer wird es, wenn nicht nur eine Variable (hier: die Größe des Unternehmens) betrachtet wird, sondern aus der Bilanz Kennzahlen abgeleitet werden, die sich aus mehreren Variablen zusammensetzen. Für Kennzahlen wie den Verschuldungsgrad oder die Eigenkapitalquote lassen sich, alle Unterpositionen einer Standardbilanz gegeben, sehr viele verschiedene Versionen definieren, und es ist a priori nicht klar, welche Version am für das Ratingmodell tauglichsten ist. Genau das heraus zu finden ist Aufgabe der Einzelfaktoranalyse.

Ist der Katalog an potenziell brauchbaren Kennzahlen definiert, muss dazu zuerst die betriebswirtschaftliche Sinnhaftigkeit einer jeden Kennzahl im Zusammenhang mit der Insolvenzrisikoversorgung getestet werden. Es wird also für den einzelnen Risikofaktor (wie

etwa den Verschuldungsgrad oder die Eigenkapitalquote) anhand des Entwicklungsdatensatzes überprüft, ob die Kennzahlenwerte für insolvente Unternehmen im Mittel eine betriebswirtschaftlich schwächere Ausprägung (d.h. z.B. geringe Eigenkapitalquote) besitzt als für solvente Unternehmen. Kennzahlen, die diese Eigenschaft nicht besitzen, werden von der weiteren Analyse ausgeschlossen.

Am einfachsten gelingt diese Analyse für Faktoren, die nur eine beschränkte Anzahl an Ausprägungen aufweisen – so, dass man Insolvenzraten für jede Ausprägung errechnen kann, deren Belegung nicht zu gering ist (mind. rd. 100 Insolvenzen pro Klasse). Für diese Art von kategorialen Variablen kann dann ein Bild erstellt werden, das jeder Klasse seine Insolvenzrate gegenüber stellt. Werden die Belegungsmengen der Klasse ebenso mit in die Grafik aufgenommen, ist das Risikoprofil des Faktors mit einem Blick zu erfassen, wie das Beispiel Branche zeigt:

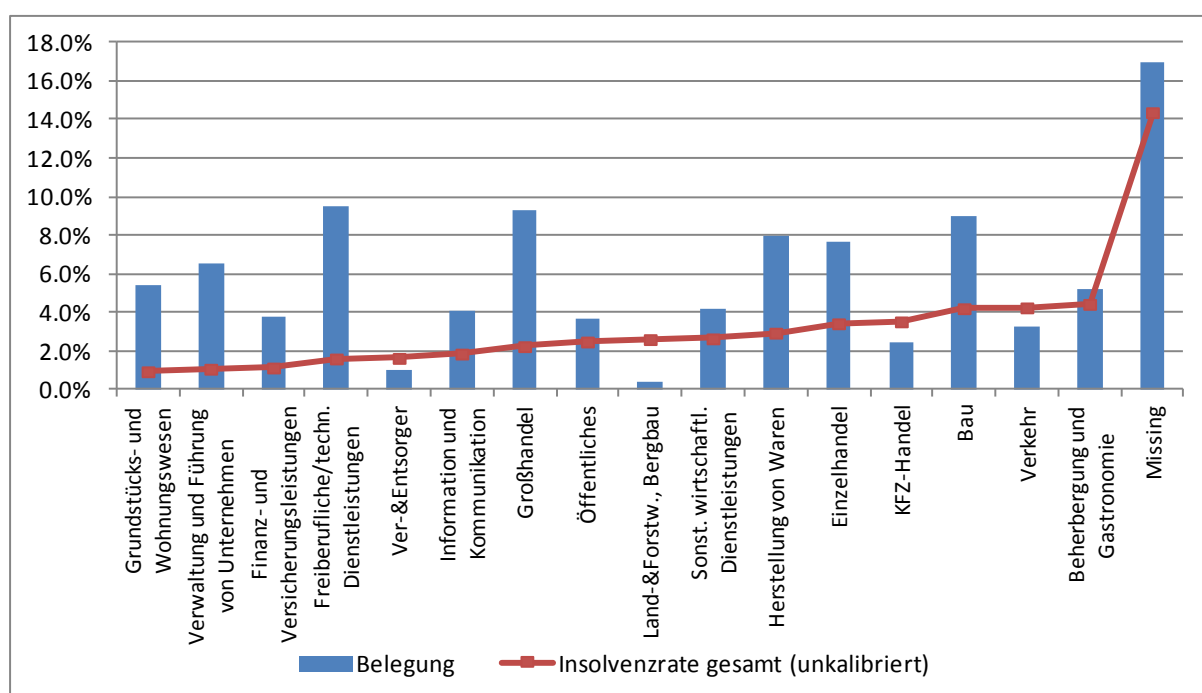


Abbildung 5: Einzelfaktoranalyse Branche

In aufsteigender Sortierung zeigt das Bild oben also, welche Branchen welches Insolvenzrisiko aufweisen (rote Linie), und wie häufig diese Branchen im gesamten Firmenbuch-Portfolio vorkommen (blaue Säulen). Würde sich hier keine Differenzierung der Insolvenzrate zeigen oder aber die Ergebnisse fachlich nicht plausibel erscheinen, wäre dieser Faktor unbrauchbar für das Ratingmodell.

Einige Variablen werden dabei erst zu brauchbaren Risikofaktoren, wenn sie eine sinnvolle Vorverarbeitung erfahren. So ist es etwa wenig hilfreich, den Sitz des Unternehmens nach seiner Postleitzahl zu bewerten, weil diese Ebene auch für das sehr große Portfolio des Firmenbuches zu granular wäre (viele PLZ enthalten <10 Unternehmen, das sind zu wenige, um eine glaubwürdige Insolvenzrate bestimmen zu können). Kategoriale Variablen dieser Art sind also vorab zu gruppieren und fachlich zu bewerten – was bei den oben gezeigten Branchen, die sich vom für das Unternehmen vorliegenden OENACE-Code ableiten, auch Freiheitsgrade übrig lässt. Sind fachlich plausible Gruppen entstanden, werde die Insolvenzraten dieser Kategorien in weights-of-evidences transformiert, eine Standardprozedur zur Einbindung kategorialer Variablen in das Regressionsverfahren.

Schwieriger wird die Verarbeitung bei Variablen, die theoretisch unendlich viele Ausprägungen besitzen. Bilanzkennzahlen wie etwa die Eigenkapitalquote sind von Natur

aus metrisch, d. h. sie können lückenlos einen ganzen Wertebereich überdecken, so dass es unmöglich wird, für jede Ausprägung eine Insolvenzrate zu errechnen (der Datensatz kennt ja z. B. nur ein einziges Unternehmen mit der Eigenkapitalquote 23,456789 %). Hierzu bedarf es einer Vorverarbeitung, die metrische Einzelfaktoren für die nachfolgende Multifaktoranalyse handhabbar macht.

Als herausragende Transformation, mittels der eine Rohkennzahl aus der Bilanz auf eine Variable gemappt werden kann, die später als Eingangsvariable in die (multivariate) Regression genutzt wird, hat sich der Hodrick-Prescott-Filter erwiesen. Der Hodrick-Prescott-Filter ist ursprünglich ein mathematisches Mittel der Makroökonomie, um Zeitreihen zu analysieren. Er wird benutzt, um eine Schwankungen auszugleichen, so dass die Zeitreihe neben dem lokalen Trend nur wenig abhängig ist von lokalen Varianzen.

Ein exemplarisches Ergebnis aus dieser Einzelfaktortransformation für metrische Variable ist in der folgenden Grafik dargestellt. Anhand der Kennzahl „Wachstum der Bilanzsumme“ kann gezeigt werden, auf welche Schwierigkeiten (Nichtnormalität, Nichtlinearität, fat tails) die Einzelfaktoranalyse stößt. Der Einzelfaktor zeigt in seiner Mengenverteilung (blaue Säulen), dass Firmen von einem Jahresabschluss auf den nächsten in ihrer Bilanzsumme ganz gehörige Sprünge machen können (Darstellung von -80 % bis +105 % Wachstum). Die Masse der Unternehmen befindet sich allerdings in einer Ab- bzw. Zunahme von -18 % bis +25 %, aber die Ausläufer der Verteilung sind massiv (fat tails).

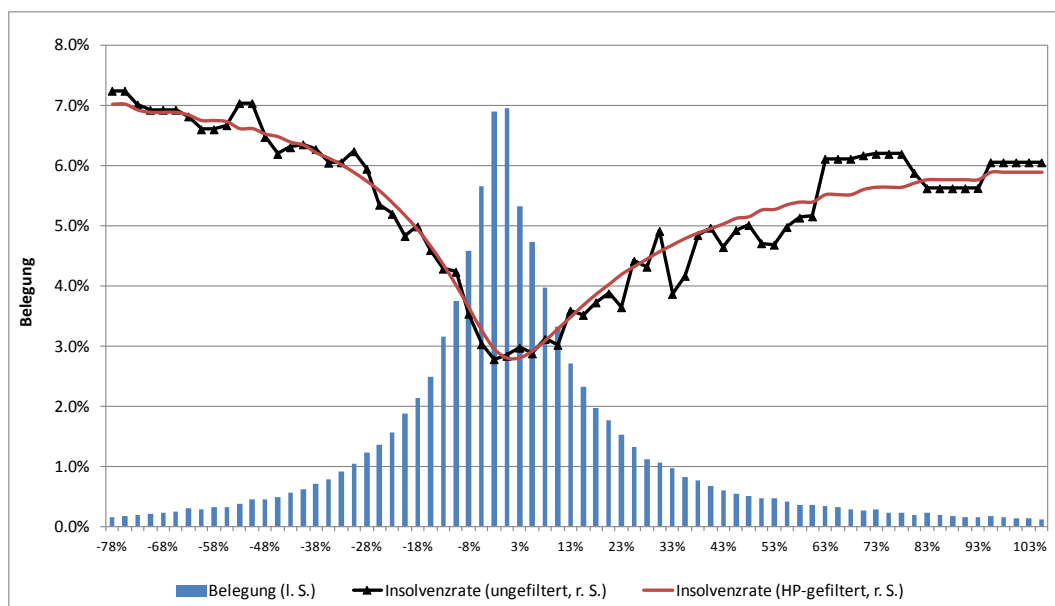


Abbildung 6: Einzelfaktoranalyse Bilanzsummenwachstum

Wird nun für einzelne Abschnitte der Wachstumsrate (etwa für alle Fälle von -48% bis -46%) eine Insolvenzrate errechnet, so ergibt sich über die ganze Skala hinweg ein uneinheitliches Bild, das nur aus der zu feinen Klasseneinteilung herrührt. Der Verlaufstrend der Insolvenzraten ist allerdings fachlich sehr plausibel: Zu großes Wachstum ist ebenso riskant wie zu großes Schrumpfen. Ein Mittelmaß von 2 % bis 5 % Wachstum scheint ideal, dort zeigen sich die geringsten Insolvenzraten. Wie die rote Linie in der Grafik oben nun zeigt, glättet der HP-Filter den in den Rohdaten sehr volatilen Verlauf (schwarze Linie). Dabei lässt sich auch erkennen, dass der HP-Filter durchaus die Nichtlinearität mitnimmt und gleich einem Tiefpassfilter nur die Struktur des Risikofaktors abgreift, nicht aber die Schwankungen darüber hinaus. Mit einem dergestalt geglätteten Risikoverlauf geht der Einzelfaktor nun also in den nächsten Modellierungsschritt, die multivariate Analyse.

## 3.4 Multifaktoranalyse

Die multivariate Analyse (Multifaktoranalyse) dient dem Ziel, die einzelnen in das Ratingmodell eingehenden Faktoren dergestalt zu gewichten, dass das Insolvenzereignis möglichst genau vorhergesagt werden kann. Das inkludiert zum einen die Auswahl der Variablen, zum anderen die gegenseitige Gewichtung im Verbund aller Kennzahlen. Während letztere mit dem Verfahren der logistischen Regression methodisch klar abgesteckt ist, ist die erstere Aufgabe, die Auswahl der Variablen, die in das Modell eingehen, über die Statistik hinaus fachlich gestaltbar.

Standardmäßig werden in der Theorie für die Variablenauswahl forward, backward oder stepwise Selektionsmechanismen vorgeschlagen – Techniken, die automatisch ablaufen, ohne dem Modellierer Freiheitsgrade zu belassen. In der Praxis ist man allerdings mit der Schwierigkeit konfrontiert, dass viele Kennzahlen, die sich zum Teil sehr ähnlich sind, zur Verfügung stehen, und daher die Wahl für eine bestimmte Kennzahl von nur marginalen Verschiedenheiten getrieben ist (Problem der Multikollinearität). Das führt dazu, dass mit den automatisch zur Verfügung stehenden Variablenselektionsverfahren schon bei leicht anderer Wahl der Entwicklungsstichprobe ein anderes Set an Variablen gezogen wird. Wenn dann nach Einsatz eines automatischen Selektionsmodus ein Rating-Modell vorliegt, lässt sich in vielen Fällen erkennen, dass die ausgewählten Kennzahlen weder ausgewogen sind, noch die Hypothesen aus der Single-Faktor-Analyse sich im multivariaten widerspiegeln (Wirkungsrichtung der Variablen wegen Korrelation verkehrt). Zudem kann festgehalten werden, dass bei genügend großer Anzahl an Kennzahlen sehr viele Modelle die praktisch gleiche (d.h.: statistisch nicht signifikant unterscheidbare) Performance aufweisen – aus ökonomischer Sicht eröffnet das die Möglichkeit, sich für eines dieser besten Modelle zu entscheiden, ohne auf statistische Güte verzichten zu müssen.

Die zur Verfügung stehenden Freiheitsgrade, ein ausgezeichnet trennscharfes Rating zu wählen, wurde im Falle des Compass-Benchmark dahingehend genutzt, dass zum einen ein Kennzahlen-Kollektiv gewählt wurde, das die Bilanz ausgewogen abdeckt, zum anderen darauf geachtet wurde, dass die aus der Einzelfaktoranalyse erkannten ökonomischen Zusammenhänge (betriebswirtschaftliche Aussage) auch im Multifaktor-Modell präsent bleiben. Darauf basierend wurde mit Hilfe einer multivariaten logistischen Regression ein Modell mit 13 Kennzahlen (10 zusammengesetzte Bilanzkennzahlen, die Bilanzsumme als Größenindikator, die Rechtsform und die Branche) entwickelt, das in allen Variablen hochsignifikant ist ( $p$ -Wert < 1 %) und zum anderen einen Sättigungsgrad aufweist, der garantiert, dass die Trennschärfe des Modells nicht durch Hinzunahme einer weiteren Kennzahl signifikant trennschärfer werden würde.

Die im Compass-Benchmark verwendeten Kennzahlen für die Berechnung der Insolvenzwahrscheinlichkeit geben Aufschluss über die Bereiche Stamminformation, Aktiv- und Passivstruktur, Ertragsstruktur, sowie über das Unternehmenswachstum. Im Folgenden werden die in das Modell eingehenden Kennzahlen und deren Einfluss auf die Solvenz eines Unternehmens näher erläutert. Nachfolgende Tabelle zeigt in der Spalte Gewicht den relativen Einfluss, den die Kennzahl auf das Ratingergebnis ausübt, nebst seiner thematischen Einordnung, der Definition, und der ökonomischen Interpretation der Kennzahl.

Rang des Gewichts	Kategorie	Kennzahl	Definition	Inhaltliche Bedeutung (Wirkungsrichtung)
1	Stammdaten	Branche	Abgeleitet aus den ersten beiden Stellen des OENACE-Codes	Gruppierung in 17 Klassen, Klassiker jedes Firmenratings.
2	Aktivstruktur	Sachanlagenintensität	Sachanlagevermögen / Bilanzsumme	Anteil der Sachanlagen an der Bilanzsumme. Eine zu geringe Sachanlagenintensität kann negativ sein, da es bedeuten kann, dass das Unternehmen überwiegend mit alten, weitgehend abgeschriebenen Anlagen arbeitet – was oftmals ein Indikator für geringe Innovationskraft ist. Hohe Sachanlageintensitäten korrespondieren zu geringer Insolvenzrate.
3	Stammdaten	Anzahl Mitarbeiter	Anzahl Mitarbeiter	Die Mitarbeiteranzahl ist ein Maß für die Größe des Unternehmens. In aller Regel weisen größere Unternehmen geringere Insolvenzhäufigkeiten auf als kleinere.
4	Aktiv-Passivstruktur	Liquidität	Liquide Mittel (Kassenbestand + Wertpapiere) / Verbindlichkeiten	Solvente Unternehmen besitzen typischerweise eine ausreichende, aber nicht zu hohe Liquidität. Eine zu hohe Liquidität lässt auf mangelnde Planung und einen nicht effizienten Mitteleinsatz schließen. Eine zu geringe Liquidität gefährdet Zahlungsziele und deutet eher auf Insolvenz hin.
5	Passivstruktur	Inverser Verschuldungsgrad	Eigenkapital / Verbindlichkeiten	Diese Kennzahl beleuchtet nur die Passiva, also die Finanzierungsseite des Unternehmens. Je weniger „gehebelt“ das Unternehmen ist, desto mehr schützt das vorhandene Eigenkapital den Gläubiger vor dem Insolvenzfall. Statistisch wirkt die Kennzahl wie eine Eigenkapitalquote, mit dem Unterschied, dass bei letzterer der Nenner aus allen Passiva besteht, während beim (inversen) Verschuldungsgrad das Eigenkapital nur mit den Verbindlichkeiten in Bezug gesetzt wird. Aus technischen Gründen wird der Verschuldungsgrad, üblicherweise definiert als Verbindlichkeiten / Eigenkapital, hier im Kehrwert gebildet, und trägt daher den Beinamen „invers“.
6	Unternehmenswachstum	Bilanzsummensteigerung	Bilanzsumme Vorjahr / Bilanzsumme akt. Jahr - 1	Wachstumsraten moderat über 0 % sind optimal, Schrumpfung sowie zu schnelles Wachstum zunehmend riskanter (s. Abbildung 6 oben).
7	Stammdaten	Bundesland	Mapping der PLZ des Unternehmens auf das Bundesland	Die österreichischen Bundesländer weisen stark unterschiedliche Insolvenzquoten auf. So zeigen die Analysen, dass die drei westlichen Bundesländer Vorarlberg, Tirol und Salzburg klar besser firmieren als die östlichen Bundesländer. Das Bundesland mit den höchsten Ausfallraten ist Kärnten.
8	Stammdaten	Rechtsform	Abgeleitet aus einer Liste an möglichen Rechtsformen	3 Ausprägungen: GmbH, AG, KG & Rest. Klassiker in Bilanzratings, weil Kapitalgesellschaften höheres Insolvenzrisiko aufweisen als Personengesellschaften.
9	Ertragsstruktur, Passivstruktur	Gewinnwachstum	(Bilanzgewinn akt. Jahr – Bilanzgewinn Vorjahr) /  Bilanzgewinn Vorjahr	Die Kennzahl zeigt die Steigerung des Bilanzgewinns des Unternehmens. Basis ist der Bilanzgewinn des Vorjahres (absolut). Je höher die Kennzahl, desto weniger Risiko, wobei ungewöhnlich hohe Gewinnwachstumsraten auch wieder tendenziell auf höhere Ausfallraten hinweisen.

10	Stamm- daten	Unter- nehmens- alter	(Stichtag - Gründungsdatum) / 365	Das Alter des Unternehmens ist ein fixer Bestandteil in praktisch jedem Ratingmodell: Je älter, desto weniger wahrscheinlich ist die Insolvenz.
11	Stamm- daten	Bilanz- summe	Summe der Aktiva = Summe der Passiva	Als Absolutgröße zeigt diese Kennzahl den Zusammenhang, dass größere Firmen weniger oft von Insolvenz betroffen sind als kleinere.
12	Ertrags- struktur, Passiv- struktur	Eigen- kapital- rendite	$(\text{Bilanzgewinn akt. Jahr} - \text{Bilanzgewinn Vorjahr}) /  \text{Eigenkapital akt. Jahr} + \text{Eigenkapital Vorjahr}  \times 2$	Die Steigerung des Bilanzgewinns wird hier in Bezug gesetzt zum mittleren Eigenkapital der letzten beiden Bilanzen. So ist es möglich, ein Maß dafür zu finden, welcher Bilanzgewinn mit dem eingesetzten Eigenkapital zu erzielen war. Je größer diese Eigenkapitalrendite, desto weniger häufig finden sich Insolvenzfälle.

Tabelle 1: Definition und Interpretation der Kennzahlen

Die Angabe des Ranges, mit welchem Gewicht der Einzelfaktor in das Gesamtmodell eingeht, ist nur indikativ: Das Gewicht kann nicht als prozentualer Effekt dargestellt werden, weil die Berechnung der Insolvenzwahrscheinlichkeit nicht einer linearen Gleichung folgt. Für jeden einzelnen Fall kann die individuelle Gewichtung daher leicht anders sein. Die Rangfolge ist aber aus der Bedeutung der Gewichte für das Gesamtportfolio erkennbar, und sie gibt einen Eindruck, welche Kennzahlen bedeutend, nur mittelmäßig wichtig oder eher unwichtig sind. Es muss beim Lesen der Tabelle allerdings stets beachtet werden, dass die Kennzahlen im Verbund wirken: Die unter 12 genannte Eigenkapitalrendite etwa wirkt für sich genommen als hervorragender Prädiktor für Insolvenz, als Einzelfaktor ist diese Kennzahl ganz stark. Werden davor aber schon Kennzahlen wie der inverse Verschuldungsgrad (5) und das Gewinnwachstum (9) betrachtet, kann die Eigenkapitalrendite kaum mehr etwas substantiell Neues hinzufügen – ihre zusätzliche Erklärungskraft liegt also zurück, weil der Großteil der Varianz schon von den vorher schlagenden Faktoren erklärt wurde.

Zur Verarbeitung der Kennzahlen in einem geschlossenen Modell werden die metrischen Variablen (also alle Faktoren bis auf die Branche und die Rechtsform) durch die Hodrick-Prescott-Transformation geschoben, und die kategorialen Faktoren (Branche, Bundesland und Rechtsform) nach fachlich sinnvollen Gruppierungen durch ihre weights-of-evidence (eine besser zu verarbeitende Abwandlung der Insolvenzrate) repräsentiert. Mit diesen Werten wird eine logistische Regression angestoßen, deren abhängige Variable der Event „Insolvenz“ ist. Resultat dieser logistischen Regression ist für jedes Merkmal ein Gewicht, das über eine analytische Formel die Wahrscheinlichkeit beschreiben lässt, dass es sich im vorliegenden Fall um einen Insolvenzfall handelt.

### 3.5 Kalibrierung auf eine mittlere Insolvenzwahrscheinlichkeit

Als abschließender Schritt in der Modellentwicklung muss die aus der logistischen Regression geschätzte Insolvenzwahrscheinlichkeit auf eine repräsentative 1-jährige Insolvenzwahrscheinlichkeit gemappt werden. Dies ist notwendig, da die Rate an Insolvenzfällen in der Entwicklungsstichprobe nicht der mittleren Insolvenzwahrscheinlichkeit der österreichischen Unternehmen entspricht – schon alleine deshalb nicht, weil in die multivariate Analyse alle Insolvenzen (rd. 4,8 %) eingehen, die finale Ratingzahl aber auf den Zeithorizont 1 Jahr Bezug nehmen sollte (pro Jahr gehen weniger als 2 % der Unternehmen in Insolvenz, s.u.). Für die Berechnung der finalen PD (probability of default, Ausfallwahrscheinlichkeit) wurde eine analytische Mapping-Formel verwendet, die Skalen verschiedener Insolvenzraten ineinander überführen lässt (s. Leitfaden OeNB).

Für die Kalibrierung des Compass-Benchmark wurde ebenso auf die langjährige Historie der Firmenbuchdaten zurückgegriffen. Mithilfe der Verweildauer der Unternehmen von Erstmeldung bis zum Stichtag des Datenabzugs (für solvente Firmen) bzw. bis zum Insolvenzdatum (für insolvente Firmen), kann abgeschätzt werden, wie hoch die Wahrscheinlichkeit ist, dass ein beliebiges Unternehmen in einem Jahr Insolvenz anmelden muss. Diese mittlere Insolvenzrate (sog. Zentraltendenz oder Anker-PD) beträgt laut Analyse im Durchschnitt der Jahre 2004-2012 für Firmenbuch-bilanzpflichtige Unternehmen 1,37 %. Für Kapitalgesellschaften liegt diese Rate leicht höher, für Personengesellschaften deutlich tiefer – aber der Schnitt aller österreichischer Unternehmen hält sich stabil unter der 2 %-Marke.

Liefert die logistische Regression also die Gewichtung der Einzelfaktoren, und die Anker-PD die mittlere Insolvenzwahrscheinlichkeit, lässt sich für jedes Unternehmen seine spezifische, kalibrierte Insolvenzwahrscheinlichkeit errechnen. Nun ist dieses Maß allerdings sperrig in der Kommunikation, sodass sich im modernen Risikomanagement einfacher zu transportierende Notensysteme eingebürgert haben. Die bekanntesten sogenannten "Masterskalen" stammen von den großen Ratingagenturen Standard&Poor's und Moody's. Gleich ist ihnen allen, dass sie nur Klassenbildungen von Insolvenzwahrscheinlichkeiten sind, denen ein sprechender Name verliehen wurde.

Für den Compass-Benchmark wird eine eigene Masterskala in Stand gesetzt, die sich am österreichischen Schulnotensystem orientiert. Zur Definition der Schnittwerte für die PD's (Insolvenzwahrscheinlichkeiten, die die einzelnen Ratingnoten trennen) wurde festgelegt, dass jede Note 10 % der Fälle aufnehmen sollte. So ist es bei 10 Noten ganz einfach, sich die Stellung gegenüber allen anderen Unternehmen zu errechnen, weil die Noten den Dezilen der Ratingverteilung entsprechen. In Anlehnung an Schulnoten von 1 bis 4- (5 würde schon Insolvenz bedeuten), definieren wir in untenstehender Tabelle:

Ratingnote	Mittlere Insolvenz-wahrsch.	Obergrenze Insolvenz-w.
1	0,050 %	0,063 %
2+	0,085 %	0,115 %
2	0,147 %	0,188 %
2-	0,235 %	0,293 %
3+	0,361 %	0,446 %
3	0,553 %	0,685 %
3-	0,865 %	1,092 %
4+	1,426 %	1,861 %
4	2,679 %	3,857 %
4-	9,820 %	99,999 %
5	100,000 %	100,000 %

Tabelle 2: Die Compass-Benchmark-Masterskala

Neben der Definition der Ratingnoten durch eine maximale Insolvenzwahrscheinlichkeit gibt es auch die Möglichkeit, die Compass-Benchmark PD via Mapping auf die Standard&Poor's-Skala zu bringen, um international vergleichbare Bonitätseinstufungen zu erhalten. Gleich ist beiden Skalen, dass die Insolvenzwahrscheinlichkeiten nach hinten hin nicht linear, sondern exponentiell ansteigen. Diese Eigenschaft macht Sinn für die genauere Differenzierung der verschiedensten Unternehmen auf einer einheitlichen Skala. Unterschiedlich ist die Belegung der einzelnen Klassen, welche bei der Compass-Skala gleichverteilt ist, bei den S&P-Skala eher zentriert, wie nachstehende Abbildung zeigt.

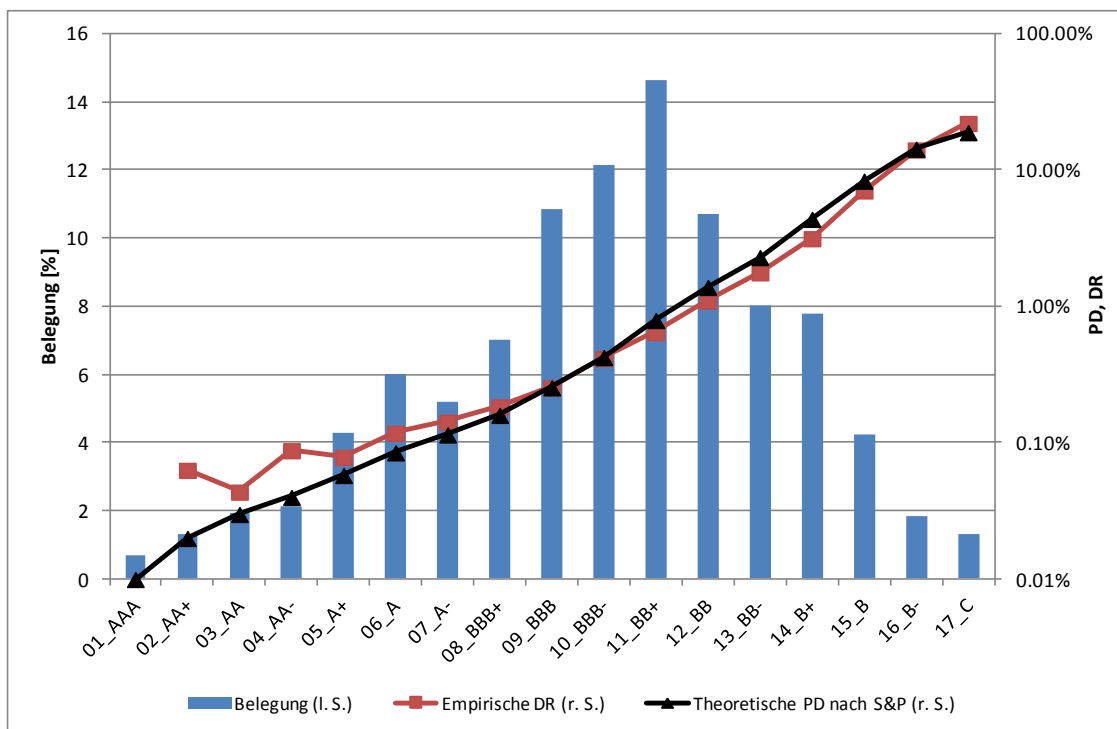


Abbildung 7: Verteilung des Gesamtportfolios über die Ratingnote

Die blauen Säulen (Belegung, linke Skala) zeigen dabei den Mengenanteil, der sich in dieser Klasse wiederfindet (also gut 4 % der Firmen erhalten ein A+), die schwarze Linie (PD, rechte, logarithmische Skala) weist die in dieser Klasse vorherrschende Insolvenzwahrscheinlichkeit aus.

Für beide Ratingskalen sei nochmals klargestellt, dass der Compass-Benchmark keine 5 bzw. kein D vergibt, weil das bedeuten würde, dass die Insolvenz schon feststeht. Sowie ein Unternehmen in Insolvenz geht, bedarf es aber keiner Prognose mehr, dass dieses Ereignis eintritt, die damit verbundene Wahrscheinlichkeit ist demnach 100 %.



## 4 Validierung

Die Validierung ist wichtige Voraussetzung für die Qualität eines Ratingmodells. So wird diese auch von der Bankenaufsicht für die Genehmigung von Basel II-fähigen Ratingsystemen bei Banken verlangt (s. Leitfaden OeNB). Wichtige Bereiche, die innerhalb der Validierung betrachtet werden, sind die Trennschärfe, die Stabilität und die Kalibrierung eines Ratingmodells.

### 4.1 Trennschärfe

Die Trennschärfe gibt an, inwieweit ein Ratingsystem zwischen „guten“ und „schlechten“ Fällen unterscheiden kann. „Gut“ bedeutet in diesem Zusammenhang den Nichteintritt der Insolvenz, „schlecht“ das Eintreten der Insolvenz innerhalb des Prognosezeitraums. Dieser Prognosezeitraum beträgt üblicherweise ein Jahr. Das am häufigsten verwendete Trennschärfemaß ist der Gini-Koeffizient, der anhand der Gini-Kurve (auch Cumulative Accuracy Profile) berechnet wird. Für diese Berechnung wurde ein Validierungssample gebildet, das unabhängig vom Entwicklungssample ist (Out-of-sample Test).

Für die Gini-Kurve werden die kumulierten Häufigkeiten für die „schlechten“ Fälle auf der y-Achse und die kumulierten Häufigkeiten aller Fälle auf der x-Achse abgetragen. Die Darstellung liest sich wie folgt: Unter den x % schlechtest-gerateten Fällen des Portfolios befinden sich y % der tatsächlich über einen 12-Monats-Horizont ausgefallenen Fälle. Basierend auf dieser Darstellung wird der Gini-Koeffizient berechnet, der die Krümmung der Gini-Kurve komprimiert in einer Kennzahl ausdrückt. Der Gini-Koeffizient setzt die Fläche zwischen dem Zufallsmodell und der Gini-Kurve des Ratingmodelles in das Verhältnis zu der Fläche zwischen Zufallsmodell und dem theoretischen perfekten Modell. Je höher der Gini-Koeffizient, desto Trennschärfer das Modell.

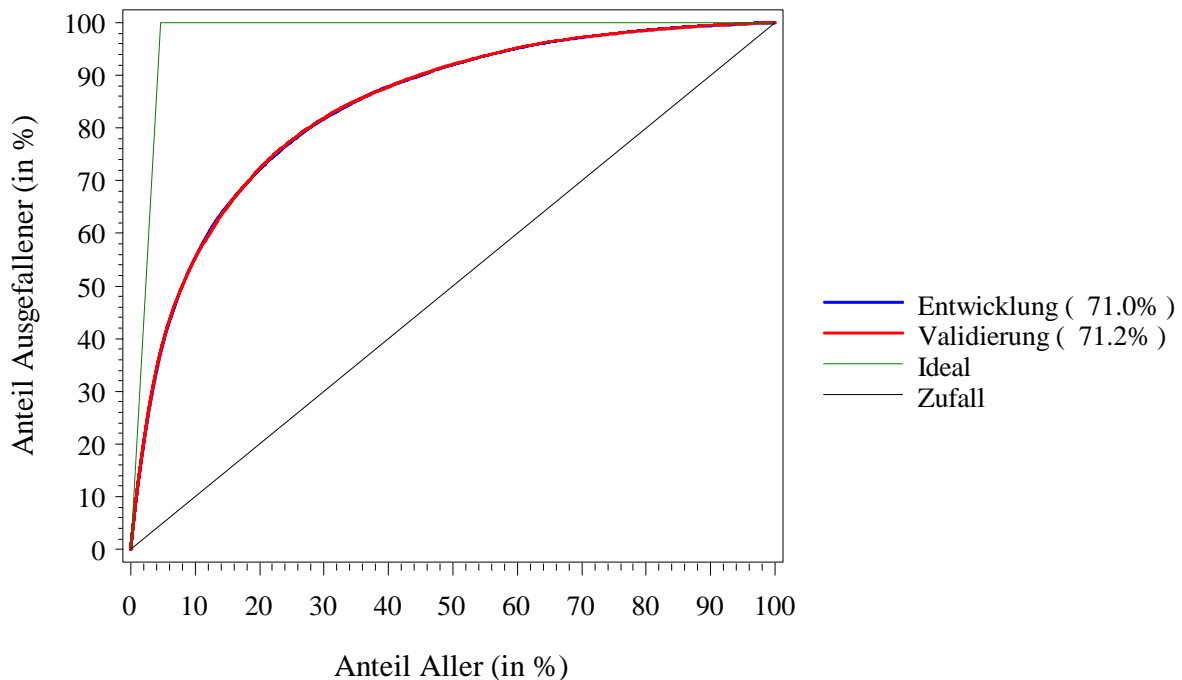


Abbildung 8: Gini-Kurve des Compass-Benchmark

Mit einem Gini-Koeffizienten von 71,2 % weist der Compass-Benchmark eine äußerst hohe Trennschärfe auf. Für Bilanzratings, die ohne Zusatzinformationen wie Frühwarnindikatoren, Zahlungsverkehrsdaten oder Anhangangaben auskommen, ist diese Trennschärfe als absolut erstklassig zu verstehen. Vergleichbare Ratingmodelle von Banken erreichen out-of-sample Gini-Koeffizienten von (nur) 50 %-65 %, sind aber in vielen Fällen mit mehr Bilanzpositionen (insbesondere GuV-Positionen) versorgt als der Compass-Benchmark. Grund für die ausgezeichnete Güte des Modells ist die sehr große Anzahl an homogen erfassten Bilanzen, die eine unschätzbar gute Datenbasis liefern.

## 4.2 Stabilität

Unter Stabilität versteht man die Eigenschaft, dass sich ein Modell nicht nur auf einem bestimmten Datenbestand (auf einem spezifischen Portfolio) bewährt, sondern dass sich über verschiedene Anwendungsfelder und Zeiträume hinweg die Struktur des Modells (etwa: die Auswahl der Variablen und die Verteilung der Ausprägungen) erhalten bleibt. Der Compass-Benchmark wurde in den verschiedensten Schnitten getestet und optimiert: Über mehrere Jahresscheiben, über mehrere Rechtsformen, über mehrere Unternehmens-Größenklassen, über mehrere geografische Regionen, und über mehrere Branchen hinweg zeigte sich, dass sich bei dessen Einsatz in neuer Umgebung noch immer ein trennscharfes Schätzprofil ergibt.

Als primärer Test für die Stabilität gilt der Vergleich der Ratingverteilungen für das Entwicklungs- und das Validierungssample. Weil das wichtigste Resultat aus dem Benchmark die Ratingnote ist, sollte erwartet werden, dass die Mengenverteilung der Noten sich von einer Zufallsstichprobe zur nächsten nicht verändert – dass das Modell also über ein Portfolio hinweg stabile Verteilungen liefert. Untenstehende Grafik weist die Verteilungen der beiden Stichproben nebeneinander aus, sodass schon optisch klar wird, dass das Modell nur sehr geringe Abweichungen von Entwicklung zu Validierung generiert. Das kann auch mit statistischen Maßzahlen (etwa: PSI, population stability index) unterstrichen werden, sowie sich die Stabilität auch auf jedem unabhängigen neuen Datenbestand nachweisen lassen kann.

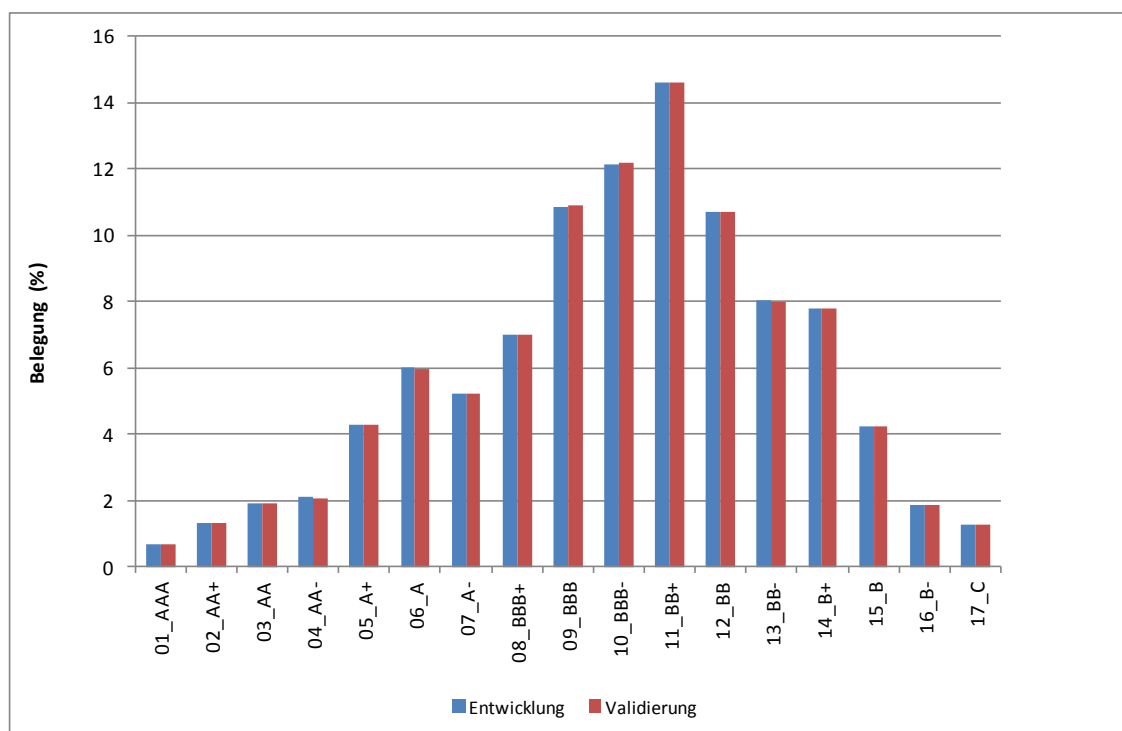


Abbildung 9: Stabilität der Ratingnote

## 4.3 Kalibrierung

Unter Kalibrierung versteht man den Zusammenhang von (ex ante) geschätzten Insolvenzwahrscheinlichkeiten (PD, probability of default) zu (ex post) tatsächlich eingetretenen Insolvenzen (DR, default rates). Kalibrierungstests überprüfen somit das Herzstück eines Ratingmodells: Die Genauigkeit der Prognose, die ein Rating für die Zahlungsfähigkeit eines Unternehmens trifft.

Die Kalibrierung lässt sich hierbei für das Gesamtportfolio messen, aber auch für einzelne Ratingklassen oder Teilssegmente an Unternehmen. Immer muss dabei die statistische Natur der Ratingnoten im Auge behalten werden: Eine Insolvenzwahrscheinlichkeit von 1 % für ein Unternehmen etwa bedeutet, dass von 100 gleichartigen Unternehmen eines im Laufe des nächsten Jahres Insolvenz anmelden wird. Ob das genau für das betrachtete Unternehmen so kommen wird, lässt die Wahrscheinlichkeitsbetrachtung selbstverständlich offen.

Kalibrierungstests funktionieren daher nur auf Portfolioteilen, die so viele Fälle umfassen, dass den prognostizierten Insolvenzfällen auch genügend reale Insolvenzfälle gegenüber stehen. Für das Firmenbuch-Portfolio ist das definitiv der Fall, sodass hier auch auf Detailebene ein genauer Vergleich von geschätzten zu beobachteten Insolvenzen erfolgen kann. Stellt man etwa auf Ebene der feinen Ratingnoten (also Schulnoten mit Differenzierung + und -) alle Unternehmen des Firmenbuches gegen die Häufigkeit an Ausfällen, die sich ein Jahr nach der Prognose eingestellt hat, so erhält man untenstehende Abbildung.

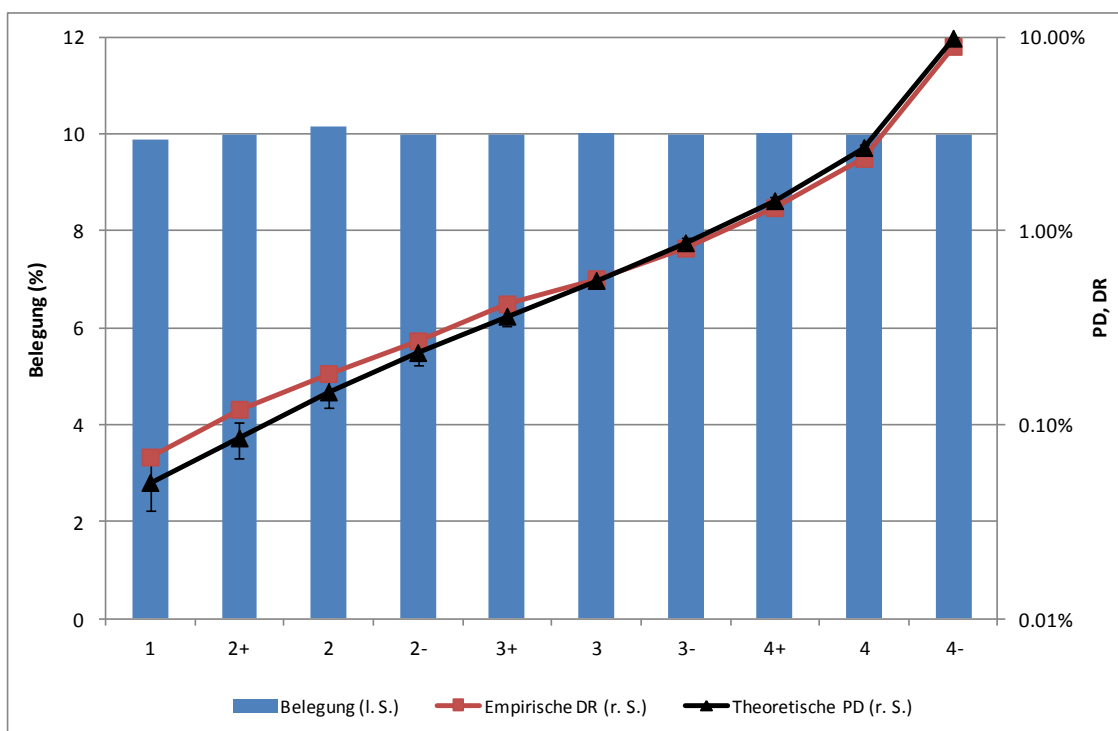


Abbildung 10: Kalibrierungstest für das Firmenbuch-Portfolio über die Ratingnoten

Die blauen Säulen markieren hierbei die Belegung – d.h. die Mengenverteilung, mit welcher sich das Gesamtportfolio in den einzelnen Ratingnoten wiederfindet (linke Skala). Alle Noten enthalten 10 % der Unternehmen, über die Schulnotenskala hinweg gibt es also eine Gleichverteilung. Die schwarze Linie weist über die Noten hinweg die prognostizierte Insolvenzwahrscheinlichkeit aus (rechte, logarithmische Skala). Unternehmen mit Ratingnote 3- wird etwa prognostiziert, zu 0,86 % auszufallen. Die rote Linie zeigt dagegen die in der nachträglichen Validierung tatsächlich aufgetretenen Ausfallraten an, d.h. die Unternehmen jeder Ratingnote, die ein Jahr nach der Notenvergabe in Insolvenz gingen (in Relation zur

Gesamtmenge der Unternehmen dieser Ratingnote; rechte Skala). In der Klasse "Ratingnote 3-" etwa fanden sich nach einem Jahr tatsächlich 0,81 % an Insolvenzfällen wieder.

Die Kalibrierungsdarstellung zeigt, dass über alle Klassen hinweg die Prognose (schwarze Linie) nahe an der Realität (rote Linie) liegt nicht. Wie optisch erkennbar, bleiben die tatsächlichen Insolvenzraten in fast allen Ratingnoten-Klassen im 95 %-Signifikanzband der theoretisch vorhergesagten Werte. Auch nach strengen statistischen Tests (etwa: Binomialtest, Hosmer-Lemeshow-Test) ist die Kalibrierung des Compass-Benchmark als sehr gut einzustufen, ein Befund, der die hohe Güte des Gesamtmodells widerspiegelt.

Validierungstests der obigen Art wird es für den Compass-Benchmark in jährlichen Abständen geben. Diese Übung dient dazu, die Güte des Modells laufend zu überprüfen – eine Vorgabe, die auch Banken für ihre Ratingmodelle durch die regulatorische Aufsicht abnehmen lassen müssen. Folgt aus der Validierung, dass das Modell signifikant zu verbessern wäre, indem man etwa neue Faktoren einbaut, diese anders gewichtet, oder aber auch nur eine andere Segmentierung der Unternehmen vornimmt, spricht man von einer Rekalibrierung. Die Rekalibrierung erneuert das Ratingmodell, indem es die aktuellsten, breitesten (d. h. am meisten Unternehmen umfassendsten) Daten einbaut und somit wiederum Trennschärfe, Stabilität und Kalibrierung optimal justiert.

## 5 Schlussbemerkung

Für die Entwicklung des Compass-Benchmark wurde eine Standardmethodik, die bei der Entwicklung von Ratingsystemen bei Banken Anwendung findet, eingesetzt. Somit wird gewährleistet, dass die Ergebnisse des Compass-Benchmark mit den Ratingergebnissen führender Banken vergleichbar sind. Ein methodisch exzellentes Arbeiten ist wesentliche Voraussetzung für die Erstellung eines sehr guten Modells. In die Entwicklung des Compass-Benchmark floss die jahrelange Erfahrung eines Entwicklerteams ein, Experten in der Entwicklung von Basel-II-fähigen Ratingsystem für Banken.

Es wurde Wert darauf gelegt, dass die Kennzahlen, die das Modell verwendet, für den Anwender verständlich sind, und dass die Kennzahlen für jedes Unternehmen aus dem Jahresabschluss nach dem Firmenbuch-Bilanzschema entnommen werden können. Dies ist wichtig, damit das Rating für das eigene Unternehmen ohne hohen Mehraufwand durchgeführt werden kann, und damit auch fremde Unternehmen einem Rating unterzogen werden können, ohne dass dies an mangelnder Datenverfügbarkeit scheitert. Die Ratingformel ist also so ausgelegt, dass sie auf jedes Unternehmen angewendet werden kann, das im Firmenbuch eine Bilanz hinterlegt. Darüber hinaus wurden nur solche Kennzahlen verwendet, die speziell für österreichische Unternehmen eine hohe Vorhersagekraft bezüglich der Insolvenzwahrscheinlichkeit besitzen.

Der Compass-Benchmark ist ein best-practice Modell für die Messung der Insolvenzwahrscheinlichkeit österreichischer, im Firmenbuch eingetragener Unternehmen. Der erweiterte Report, in den der Compass-Benchmark eingebettet ist, bietet eine Vielzahl weiterer Vorteile, die aus diesem White Paper nicht hervorgehen. So bietet der umfangreiche Bericht etwa eine Analyse jedes einzelnen Risikofaktors für die Bilanzen des betrachteten Unternehmens im Vergleich zu seinem Branchen-Benchmark. Dadurch wird eine Offenlegung des Modells erreicht, die bei Banken, welche das Thema Rating eher als Black-Box verstehen, kaum zu erzielen ist.

## **6 Anhang**

### **6.1 Literatur**

Leitfadenreihe zum Kreditrisiko, Ratingmodelle und -validierung, OeNB & FMA, Wien, 2004.

### **6.2 Abbildungsverzeichnis**

Abbildung 1: Zeitliche Staffelung der Bilanzen.....	4
Abbildung 2: Bilanzschema mit HGB-Kennungen.....	5
Abbildung 3: Mengengerüst des Gesamtdatenbestandes.....	5
Abbildung 4: Entwicklungsprozess Compass-Benchmark .....	7
Abbildung 5: Einzelfaktoranalyse Branche.....	10
Abbildung 6: Einzelfaktoranalyse Bilanzsummenwachstum.....	11
Abbildung 7: Verteilung des Gesamtportfolios über die Ratingnote .....	16
Abbildung 8: Gini-Kurve des Compass-Benchmark .....	17
Abbildung 9: Stabilität der Ratingnote.....	18
Abbildung 10: Kalibrierungstest für das Firmenbuch-Portfolio über die Ratingnoten.....	19

### **6.3 Tabellenverzeichnis**

Tabelle 1: Definition und Interpretation der Kennzahlen .....	14
Tabelle 2: Die Compass-Benchmark-Masterskala .....	15