

DIPLOMARBEIT

Titel der Diplomarbeit

„Erstellung eines Kreditrisikomodells mittels Logit-
Regression für die österreichische Hotel- und
Tourismusbranche“

Verfasserin

Judith Spiegl

angestrebter akademischer Grad

Magistra der Sozial- und Wirtschaftswissenschaften
(Mag. rer. soc. oec.)

Wien, 2012

Studienkennzahl lt. Studienblatt:
Studienrichtung lt. Studienblatt:
Betreuer / Betreuerin:

A 157
Diplomstudium Internationale Betriebswirtschaft
Univ.-Prof. Dr. Alexander Stomper

Inhaltsverzeichnis

Abbildungsverzeichnis	5
Tabellenverzeichnis	7
Symbolverzeichnis	9
Kapitel 1. Einleitung	13
Kapitel 2. Kreditrisiko	17
1. Kreditrisiko: Definition	17
2. Arten des Kreditrisikos	19
3. Kreditrisikomodelle	19
3.1. Expertensysteme	21
3.2. Diskriminanzanalyse	22
3.3. Das lineare Wahrscheinlichkeitsmodell	24
3.4. Das Probit-Modell	26
3.5. Das Logit-Modell	28
3.6. Neuronale Netzwerke	30
3.7. Hybride Systeme	33
4. <i>Accuracy Ratio</i>	33
Kapitel 3. Die österreichische Hotelindustrie	39
Kapitel 4. Kennzahldefinition und Datenerhebung	45
1. Definition der Kennzahlen	45
1.1. Überblick zur Klassifizierung von Kennzahlen	45
1.2. <i>Leverage</i>	47
1.3. <i>Debt Coverage</i>	48
1.4. <i>Liquidity</i>	48
1.5. <i>Activity Ratios</i>	50
1.6. <i>Productivity</i>	50
1.7. <i>Turnover</i>	51

1.8. <i>Profitability</i>	51
1.9. <i>Size</i>	51
1.10. <i>Growth Rates</i>	52
1.11. <i>Change of Leverage</i>	52
2. Daten und beschreibende Statistik	52
Kapitel 5. Analyse	57
1. Ausfall	57
2. Selektion der Kennzahlen	57
2.1. Leverage	60
2.2. Debt Coverage	61
2.3. Liquidity	61
2.4. Activity Ratios	62
2.5. Productivity	62
2.6. Turnover	62
2.7. Profitability	63
2.8. Size	63
2.9. Growth Rates	63
2.10. Change of Leverage	64
3. Herleitung des Multivariaten Logitmodells	64
Kapitel 6. Zusammenfassung	67
Literaturverzeichnis	69

Abbildungsverzeichnis

2.1	Wahrscheinlichkeitsverteilung der Tilgungszahlungen risikobehafteter Aussenstände	18
2.2	Schematische Darstellung eines neuronalen Netzwerks	31
2.3	Aufbau eines aktiven Neurons	32
2.4	Cumulative Accuracy Profiles	36
3.1	Übernachtungen nach Unterkunftsart	42

Tabellenverzeichnis

2.1	Modellannahmen der untersuchten binären <i>response</i> -Modelle	30
3.1	Die wichtigsten Tourismusdestinationen 2008	40
3.2	Tourismus in der europäischen Union	41
4.1	Übersicht der Beobachtungen	53
4.2	Mittelwerte und Standardabweichungen der erhobenen Kennzahlen	55
5.1	Klassifizierung der Kennzahlen und Hypothesenbildung	59
5.2	Ergebnisse der univariaten Logit-Regression: Leverage	60
5.3	Ergebnisse der univariaten Logit-Regression: Debt Coverage	61
5.4	Ergebnisse der univariaten Logit-Regression: Liquidity	61
5.5	Ergebnisse der univariaten Logit-Regression: Activity Ratios	62
5.6	Ergebnisse der univariaten Logit-Regression: Productivity	62
5.7	Ergebnisse der univariaten Logit-Regression: Turnover	62
5.8	Ergebnisse der univariaten Logit-Regression: Profitability	63
5.9	Ergebnisse der univariaten Logit-Regression: Size	63
5.10	Ergebnisse der univariaten Logit-Regression: Growth Rates	63
5.11	Ergebnisse der univariaten Logit-Regression: Change of Leverage	63
5.12	Ergebnisse der univariaten Logit-Regression: Basismodell	64
5.13	Multivariate Logit-Regression Schritt 1	64
5.14	Multivariate Logit-Regression Schritt 2	64
5.15	Multivariate Logit-Regression Schritt 3	65

Symbolverzeichnis

a_i	Schwellenwert des Neuron i
a_p	Fläche unter der CAP-Kurve des perfekten Modells
a_R	Fläche unter der CAP-Kurve des evaluierten Modells
AR	<i>accuracy ratio</i>
AUC	<i>area under the curve</i> , Fläche unter der ROC-Kurve
β	Parametervektor der unabhängigen Variable; Diskriminanzkoeffizientenvektor
C	Zahlungsmittel und -äquivalente
CAP	<i>cumulative accuracy profile</i>
E	Unternehmensergebnis (siehe Fußnote 2 auf Seite 46)
EBIT	<i>earnings before interest and taxes</i>
EK	Eigenkapital
ϵ_i	Schätzfehler der Beobachtung i
$f(\cdot)$	Aktivierungsfunktion des Neuron
$F(\cdot)$	<i>linking</i> -Funktion
FK	Fremdkapital
FK^t	Fremdkapital in der laufende Periode t
FK^{t-1}	Fremdkapital in der vorangegangenen Periode $t - 1$
FK_B	Bankverbindlichkeiten
FK_B^t	Bankverbindlichkeiten in der laufende Periode t
FK_B^{t-1}	Bankverbindlichkeiten in der vorangegangenen Periode $t - 1$
FK_K	Kurzfristige Verbindlichkeiten
FK_L	Langfristige Verbindlichkeiten
k_1	FK/V , Kennzahl der Klasse <i>leverage</i>
k_2	EK/V , Kennzahl der Klasse <i>leverage</i>
k_3	FK_L/V , Kennzahl der Klasse <i>leverage</i>
k_4	FK_B/V , Kennzahl der Klasse <i>leverage</i>
k_5	FK_B/FK , Kennzahl der Klasse <i>leverage</i>
k_6	$EBIT/z$, Kennzahl der Klasse <i>debt coverage</i>

k_7	V_K/FK_K , Kennzahl der Klasse <i>liquidity</i>
k_8	V_K/FK , Kennzahl der Klasse <i>liquidity</i>
k_9	$(V_K - FK_K)/V$, Kennzahl der Klasse <i>liquidity</i>
k_{10}	FK_K/V , Kennzahl der Klasse <i>liquidity</i>
k_{11}	V_K/V , Kennzahl der Klasse <i>liquidity</i>
k_{12}	$(V_K - FK_K)/U$, Kennzahl der Klasse <i>liquidity</i>
k_{13}	V_K/U , Kennzahl der Klasse <i>liquidity</i>
k_{14}	$(C + R)/U$, Kennzahl der Klasse <i>liquidity</i>
k_{15}	$(V_K - FK_K)/FK_K$, Kennzahl der Klasse <i>liquidity</i>
k_{16}	$(V_K - vor)/FK_K$, Kennzahl der Klasse <i>liquidity</i>
k_{17}	vor/U , Kennzahl der Klasse <i>activity</i>
k_{18}	vor/ma , Kennzahl der Klasse <i>activity</i>
k_{19}	R/U , Kennzahl der Klasse <i>activity</i>
k_{20}	R/FK , Kennzahl der Klasse <i>activity</i>
k_{21}	R/ma , Kennzahl der Klasse <i>activity</i>
k_{22}	R/vor , Kennzahl der Klasse <i>activity</i>
k_{23}	per/U , Kennzahl der Klasse <i>productivity</i>
k_{24}	$(U - ma)/per$, Kennzahl der Klasse <i>productivity</i>
k_{25}	U/V , Kennzahl der Klasse <i>turnover</i>
k_{26}	$EBIT/V$, Kennzahl der Klasse <i>profitability</i>
k_{27}	$EBIT/U$, Kennzahl der Klasse <i>profitability</i>
k_{28}	V/p^* , Kennzahl der Klasse <i>size</i>
k_{29}	U/p^* , Kennzahl der Klasse <i>size</i>
k_{30}	U^t/U^{t-1} , Kennzahl der Klasse <i>growth rate</i>
k_{31}	$(FK^t/V^t)/(FK^{t-1}/V^{t-1})$, Kennzahl der Klasse <i>leverage change</i>
k_{32}	$(FK_B^t/v^t)/(FK_B^{t-1}/V^{t-1})$, Kennzahl der Klasse <i>leverage change</i>
LQ_1	Liquidität 1. Grades, <i>cash ratio</i>
LQ_2	Liquidität 2. Grades, <i>quick ratio</i>
LQ_3	Liquidität 3. Grades, <i>current ratio</i>
ma	Materialaufwand
net_i	Nettoeingabewert des Neuron i
o_h	Output des Neuron h
Ω_i	Menge der Information, die über Kreditnehmer i vorliegt
p^*	Verbraucherpreisindex
$P_i(y_i = 1)$	Wahrscheinlichkeit einer Insolvenz

$P_i(y_i = 0)$	Wahrscheinlichkeit einer Solvenz, mit $P_i(y_i = 0) = 1 - P_i(y_i = 1)$
per	Personalaufwand
$\Phi(x)$	Verteilungsfunktion der Standardnormalverteilung
R	Debitorenforderungen
ROC	<i>receiver operating characteristic</i>
ROI	<i>return on investment</i>
$\Sigma(\cdot)$	Propagierungsfunktion des Neuron
U	Umsätze
U^t	Umsätze in der laufende Periode t
U^{t-1}	Umsätze in der vorangegangenen Periode $t - 1$
V	Vermögen, Bilanzsumme; Es gilt $V = EK + FK$
V^t	Vermögen in der laufende Periode t
V^{t-1}	Vermögen in der vorangegangenen Periode $t - 1$
vor	Vorräte
VS_1	<i>gearing</i>
VS_2	Anspannungsgrad
w_{hi}	Gewichtung, die Neuron i dem Output des Neuron h beimisst
X_i	Vektor relevanter Kennzahlen aus Ω_i
y^*	Diskriminanzwert, auch <i>cut-off point</i>
y_i	abhängige Binärvariable ($y_i = 0$ oder $y_i = 1$) des Kreditrisiko-modells
y_i°	latente Variable
z	Zinsaufwand
z^d	Diskriminanzfunktion des Altman <i>Z-score</i> -Modells

KAPITEL 1

Einleitung

Im Jahr 2008 verursachten 6.318 Insolvenzen in Österreich Insolvenzverbindlichkeiten in der Höhe von 2,5 Mrd. Euro. Das entspricht knapp über einem Prozent des österreichischen Bruttoinlandsprodukts. Davon wurden 3.346 Konkursanträge mangels Masse abgelehnt, was einem Totalausfall der Forderungen der Gläubiger gleichkommt.¹ Vor diesem Hintergrund kann die Bedeutung von Kreditrisikobewertungen nicht überschätzt werden.

In diesem Zusammenhang ist eine Analyse der Insolvenzursachen interessant: So sind 68% der Insolvenzen unmittelbare Folgen falscher Managemententscheidungen und werden durch Fahrlässigkeit, durch Unwissenheit und im schlimmsten Fall auch vorsätzlich herbeigeführt.² Gerade für solche Fälle, wo die Fehler und Verlustquellen im innerbetrieblichen Bereich liegen, sollten Kreditrisikomodelle, die auf internen Unternehmenskennzahlen beruhen, Aufschluss über die Liquidität eines Unternehmens geben können.

Im Zusammenspiel mit einigen "weichen" Faktoren, die eher die qualitative Parameter berücksichtigen, lassen sich damit Rating-Modelle erstellen, wie sie zum Beispiel der KSV anbietet.³ Wenngleich aber einige Kennzahlen branchenübergreifend angewandt werden können, so ist bei Branchen, die eine stark homogene Struktur aufweisen, eine Bestimmung spezifischer Kennzahlen sinnvoll. Eine solche Branche ist die österreichische Hotelindustrie, die in dieser Arbeit untersucht werden soll.

Der Tourismus ist ein wichtiger Bestandteil der österreichischen Volkswirtschaft. Im Jahr 2003 beliefen sich die direkten Wertschöpfungseffekte des

¹vgl. KSV Kantner (2009a).

²vgl. KSV Kantner (2009b).

³Der KSV bietet das Produkt "KSV-Rating" für seine Kunden als Dienstleistung an. Siehe dazu KSV Kreditschutzverband (2008).

Tourismus auf 15,10 Mrd. Euro, was einem Anteil an der Gesamtwertschöpfung von 6,7% entspricht. 50% aller Betten entfielen 2004 auf Hotelbetriebe und gewerbliche Ferienwohnungen oder -häuser,⁴ wobei Klein- und Mittelbetriebe vorherrschend sind.⁵

Im Jahr 2003 haben die gewerblichen Beherbergungsbetriebe (Hotels, Gasthöfe und Pensionen) Bruttoinvestitionen in der Höhe von 660 Millionen Euro getätigt und 90.613 Menschen im Jahresschnitt beschäftigt.⁶ Damit ist ein Kreditrisikomodell für diese Branche nicht nur aus finanzwirtschaftlicher Sicht für den Gläubiger sinnvoll, sondern auch aus volkswirtschaftlicher Sicht zur Sicherung von Arbeitsplätzen.

Diese Arbeit liegt im Schnittpunkt der beiden Themenkreise, die bereits angerissen worden sind: Kreditrisiko und Hotelindustrie. Die zentrale Fragestellung dieser Arbeit ist nun folgende: Kann für die österreichische Hotelindustrie ein Kreditrisikomodell erstellt werden, das die branchenspezifischen Besonderheiten berücksichtigt?

In Kapitel 2 werden zunächst verschiedene Kreditrisikomodelle vorgestellt. Der Methodologie Saunders (1997) folgend zeigen wir, dass zur Bestimmung des Kreditrisikos im wesentlichen die Finanzkennzahlen der Unternehmen relevant sind. In weiterer Folge gebe ich daher einen kurzen Überblick über Kennzahlensysteme, um im letzten Abschnitt des Kapitels zu untersuchen, mit welchen Methoden aus diesen Kennzahlen jene gefiltert werden können, die signifikante Aussagen über das Kreditrisiko eines Betriebes treffen.

Im anschließenden Kapitel 3 wird die österreichische Hotelindustrie beschrieben. Die Besonderheiten der Branche als relativ homogene Struktur soll dabei unterstrichen werden. Zwei Punkte scheinen für die Evaluierung von besonderer Bedeutung zu sein: So handelt es sich bei österreichischen Tourismusbetrieben um traditionelle, klein strukturierte Hotelbetriebe mit durchschnittlich 40 Betten, die in der Regel als Familienbetriebe geführt

⁴Die verbleibenden Betten werden von privat geführten Betrieben bereitgestellt (Privatquartiere und private Ferienwohnungen und -häuser) oder entfallen auf andere Beherbergungsbetriebe wie Jugendherbergen, Schutzhütten, Campingplätze und dergleichen. Diese sind für die vorliegende Studie nicht relevant.

⁵vgl. Statistik Austria Laimer und Schischeg (2009).

⁶vgl. Statistik Austria Statistik Austria (2009a).

werden. Zum anderen weisen die Betriebe einen branchenspezifisch hohen Verschuldungsgrad auf.

Kapitel 4 und 5 wenden sich schliesslich der Aufarbeitung und Analyse der Daten zu. Es wird eine Logit-Regressionsanalyse auf einen Datensatz von insgesamt 654 heimischen Betrieben, die allesamt Klein- oder Mittelbetriebe sind und einer Bilanzierungspflicht unterliegen,⁷ durchgeführt, in Anlehnung an die Arbeit von Evelyn Hayden (2002): Nach einer formalen Darstellung der Logit-Methode werden Kandidaten-Kennzahlen bestimmt. Nach dieser Vorauswahl nach ökonomisch-qualitativen Kriterien, werden mit der statistischen Methode schliesslich die Signifikanz der Auswahl bestimmt und die Ergebnisse dargestellt.

Kapitel 6 fasst die wichtigsten Punkte nochmals zusammen und schliesst die Arbeit ab.

⁷Diese Einschränkung ist notwendig, da Bilanzkennzahlen zur Anwendung kommen sollen. Die Mehrzahl der Unternehmen im Datensatz müssen kraft Rechtsform Jahresabschlüsse nach §4 HGB erstellen, da es sich um Gesellschaften mit beschränkter Haftung handelt.

KAPITEL 2

Kreditrisiko

In diesem Kapitel werden die Definitionen des Kreditrisikos erläutert, sowie Modelle zur Einschätzung des Ausfallsrisikos beschrieben.

1. Kreditrisiko: Definition

Kreditrisiko ist die Gefahr von Verlusten aufgrund von Bonitätsverschlechterungen eines Vertragspartners bis hin zum Ausfall¹.

Saunders (1997) definiert Kreditrisiko als *“the risk that the promised cash flows from loans and securities held by F[inancial]I[institution]s may not be paid in full”*. Wird das Grundkapital auf alle, den Kreditinstituten zugesagten Cash Flows, zur Fälligkeit voll einbezahlt und finden Zinszahlungen zum versprochenen Zeitpunkt statt, bekommen Kreditinstitute das verliehene Grundkapital plus den jeweiligen Zinsen zurück. Sollte nun der Kreditnehmer seinen Verpflichtungen nicht nachkommen, sind sowohl das entliehene Kapital als auch die erwarteten Zinszahlungen einem Risiko ausgesetzt. Deswegen versprechen finanzielle Verpflichtungen von Kreditnehmern gegenüber Kreditinstituten nur einen eingeschränkten bzw. begrenzten Gewinn.

Die Variabilität der Cash Flows der Unternehmen spiegelt sich in den vielfältigen Graden des Kreditrisikos wieder. Die gesamte Bandbreite ist in Abbildung 1 illustrativ dargestellt. Einerseits kann der Kreditnehmer teilweise oder gänzlich unfähig sein, den Zinszahlungen nachzukommen. Dies entspricht dem Bereich zwischen A und B in der Grafik. Es kann auch die Kapitaltilgung betroffen sein, dies ist der Bereich zwischen B und C. Betrachtet man einerseits den begrenzten Gewinn und das mögliche Risiko andererseits, ist es verständlich, dass Kreditinstitute das erwartete Kreditrisiko einschätzen und Risikoprämien verlangen.

¹Vgl. Ott (2001).

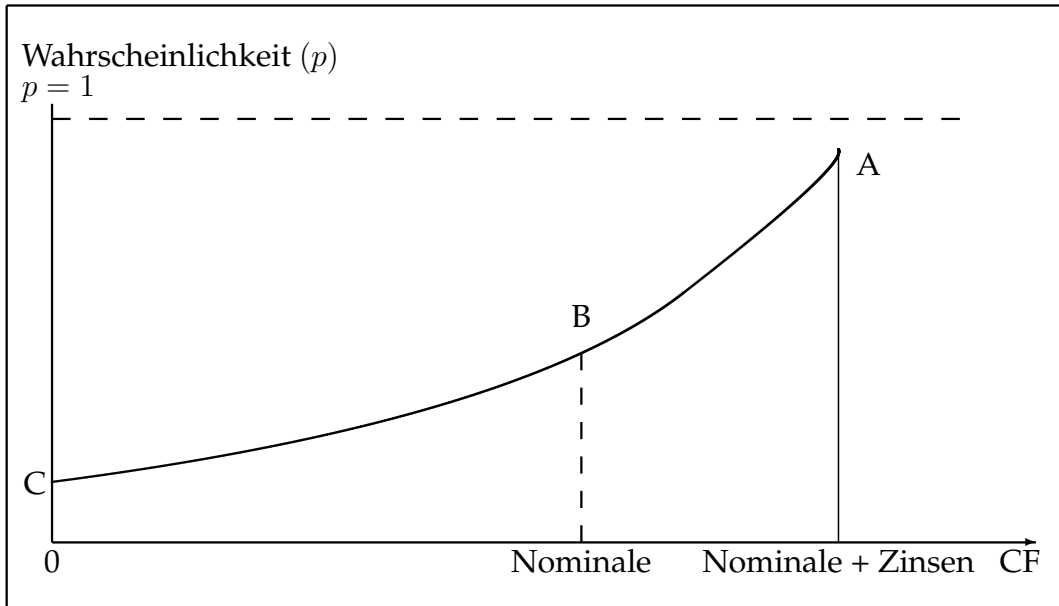


ABBILDUNG 2.1. Wahrscheinlichkeitsverteilung der Tilgungszahlungen risikobehafteter Aussenstände. Auf der horizontalen Achse sind die Tilgungszahlungen aufgezeichnet. Nominale und Zinsen werden im Punkt A zur Gänze gezahlt. Zwischen A und B ist der Schuldner beschränkt insolvent und kann das Nominale, aber nicht sämtliche Zinsen zahlen. Zwischen B und C kann der Schuldner auch das Nominale nicht zur Gänze bzw. überhaupt nicht mehr aufbringen. Quelle: Saunders (1997) S. 143.

Eine effiziente Prüfmethode sowie eine gute Kreditrisikomanagementstrategie beeinflussen Gewinn und Risiko des Portfolios direkt. Abbildung 1 zeigt die Verteilung des Kreditrisikos für einen einzelnen Posten, der dem Ausfallrisiko ausgesetzt ist. Eine Diversifikation reduziert das Kreditrisiko eines Portfolios insgesamt. Durch die Diversifikation wird allerdings nur das unternehmensspezifische Kreditrisiko verringert, das systematische Kreditrisiko bleibt erhalten. Zum systematischen Kreditrisiko gehören Faktoren, die das Ausfallrisiko sämtlicher Unternehmen in der gesamten Wirtschaft erhöhen, wie z. B. eine Rezession.

2. Arten des Kreditrisikos

Buschgen (1999) unterscheidet anhand von risikoauslösenden Sachverhalten folgende Einzelrisiken:

- **Ausfallrisiko:** das Risiko, dass der Kreditnehmer seiner Verpflichtung zur Kredittilgung sowie der Zahlung der vereinbarten Zinsen, Provisionen und Gebühren nicht oder nur zum Teil nachkommt
- **Terminrisiko:** die Gefahr einer verzögerten Zahlung durch den Schuldner, sei es, dass der Kreditnehmer den Kreditbetrag und/oder Zinsen, Provisionen, und Gebühren nicht vertragsgemäß bzw. fristgemäß entrichtet
- **Besicherungsrisiko:** die Gefahr, dass die der Bank eingeräumten zusätzlichen Sicherheiten wegen rechtlicher Mängel oder auch Wertminderungen nicht zur Deckung der Kreditforderung ausreichen
- **Geldwertrisiko:** die Bank erhält zwar den erwarteten Nominalbetrag der Kreditforderung, der durch Geldentwertung aber weniger Kaufkraft enthält

Das Bonitätsrisiko setzt sich aus dem Ausfall- und dem Terminrisiko zusammen. Das Besicherungsrisiko entsteht im Anschluss an das Bonitätsrisiko. Bonitätsrisiko und Besicherungsrisiko ergeben das aktive Kreditrisiko, das auch als Kreditrisiko im engeren Sinne bezeichnet wird. Das Kreditrisiko im weiteren Sinne umfasst neben dem Kreditrisiko im engeren Sinne auch das Liquiditätsrisiko, das Zinsänderungsrisiko und das Währungsrisiko. Es bezeichnet die Gefahr, dass der Kreditgeber die mit einem Kreditengagement verbundenen Ziele nicht im angestrebten Ausmaß realisieren kann.

3. Kreditrisikomodelle

Caouette et al. (1998) unterscheiden Modelle zur Einschätzung des Kreditrisikos anhand von drei Dimensionen: anhand des Anwendungsgebietes im Kreditverlauf, anhand der Produkte, an denen sie angewendet werden, und schliesslich anhand der angewendeten Methoden. Während die ersten beiden Dimensionen für die unterschiedlichen Anwendungsgebiete festlegen, wann ein Kreditrisiko besteht und präzise formulieren, wann ein Ausfall

eintritt, unterscheidet das letzte Klassifizierungsmerkmal die theoretischen Grundlagen, die zur Bestimmung der Höhe des Kreditrisikos herangezogen werden. Diese sollen in diesem Abschnitt untersucht werden.

Der Kategorie der qualitativen Prognoseverfahren werden Expertensysteme zugeordnet. Die ökonometrischen Prognoseverfahren umfassen im Wesentlichen die Diskriminanz- und Regressionsanalyse sowie das Probit- und Logit-Modell. Dabei soll die Ausfallswahrscheinlichkeit als abhängige Variable, deren Varianz von einer Menge von unabhängigen Variablen erklärt wird, geschätzt werden. Die unabhängigen Variablen beinhalten Kennzahlen und andere Indikatoren sowie externe Indikatoren. Saunders (1997) bezeichnet diese Systeme *credit scoring* Modelle. Damit soll der Kreditnehmer in verschiedene Kategorien bezüglich seines Ausfallsrisikos eingeteilt werden. Bei univariaten buchhaltungsbezogenen *credit scoring*-Methoden vergleicht der Analyst verschiedene Kennzahlen mit industriebezogenen Normen und Trends für jene spezifischen Variablen, um festzustellen, ob eine bestimmte Kennzahl vom Industrietrend abweicht. Die Weiterentwicklung dieses Systems führte zu den multivariaten Methoden, die weiter unten behandelt werden. Zuletzt werden neuronale Netzwerke als neuerer Ansatz sowie Hybridmodelle, die verschiedene Methoden kombinieren, untersucht.

Ganz unabhängig davon, wie der Kreditausfall definiert wird, haben alle Prognosemethoden gemeinsam, dass nur die zwei Fälle "Ausfall" und "Kein Ausfall" unterschieden werden: die prognostizierte Variable ist also binär.² Weiters soll die Wahrscheinlichkeit für diese Zustände anhand einer Informationsmenge ermittelt werden. Die Informationsmenge beinhaltet in der Regel finanzielle und operative Kennzahlen, da bereits Beaver (1968) zeigen konnte, dass bestimmte Kennzahlen zwischen gescheiterten und nichtgescheiterten Unternehmen bereits fünf Jahre vor dem Ausfall unterscheiden können. Bezeichne y_i die abhängige binäre Variable, die die Werte $y_i = 1$ als Bezeichner für den Kreditausfall und $y_i = 0$ anderenfalls annehmen kann, sowie Ω_i die Informationsmenge, die die Informationen über den Kreditnehmers i enthält, die dem Kreditgeber zur Verfügung

²Eine Erweiterung der in dieser Arbeit betrachteten logistischen Regression ist die multinomiale logistische Regression. Mit ihr ist man in der Lage, unabhängige Variable mit mehr als zwei Ausprägungskategorien zu untersuchen, vgl. auch Backhaus et al. (2008) S. 275 und Kleinbaum und Klein (2010) S. 429ff.

steht, dann entspricht die Wahrscheinlichkeit eines Ausfalls $P_i(y_i = 1)$ dem Erwartungswert $E(y_i)$:³

$$(1) \quad P_i(y_i = 1|\Omega_i) = E(y_i|\Omega_i)$$

Typischer Weise wird bei quantitativen Methoden aus der Informationsmenge Ω_i ein Vektor X_i gebildet, der aus den Werten der relevanten Kennzahlen besteht. Es werden Kennzahlen verwendet um die Ausfallswahrscheinlichkeit einzuschätzen. Dazu reicht eine kleine Auswahl aus empirisch hergeleiteten Kennzahlen für erfahrene Fachleute um präzise Aussagen über die mögliche Ausfallrate zu machen.⁴ Edmister (1972) bestätigt die Aussagekräftigkeit von Kennzahlen auch für kleine Betriebe. Die Informationsmenge Ω_i kann aber zusätzlich qualitative Informationen enthalten, die z. B. in Expertensystemen ausgewertet werden.

3.1. Expertensysteme. Anhand von spezifizierten Regeln, wird versucht die Erfahrung von Experten durch klare Entscheidungsvorgaben zu strukturieren und diese Expertise auch dem Rest des Unternehmens zukommen zu lassen. Expertensysteme gehören zu den qualitativen Modellen⁵. Vom Kreditnehmer abhängende Faktoren (Reputation, Verschuldungsgrad, Volatilität des Einkommens, Sicherheiten) sowie marktspezifische Faktoren (Konjunkturverlauf, die Höhe der Zinsen) werden von einem Experten abgewogen. Laitinen et al. (1997) zeigen, dass bei Klein- und Mittelbetrieben qualitative Indikatoren als Frühwarnsignale (*early warning sign*) besser geeignet sind als klassische finanzielle oder operative Kennzahlen. An erster Stelle dieser Indikatoren steht die mangelnde Motivation der Mitarbeiter im Rechnungswesen (*low accounting staff morale*), gefolgt von Investitionen in bzw. Abhängigkeit vom Erfolg eines großen Projekts (*invest or dependence on the success of one big project*). Durchaus wichtige Faktoren betreffen auch das Management, wie die Indifferenz des Geschäftsführers für sein Unternehmen (*Indifference of the MD about business*) oder mangelnde Entscheidungskompetenz (*poor decision-making ability of the MD*).

³Vgl. Davidson und MacKinnon (1993)

⁴Vgl. Libby (1975).

⁵Vgl. Saunders (1997), S. 275.

3.2. Diskriminanzanalyse. Die Diskriminanzanalyse⁶ ist ein strukturenprüfendes multivariates Verfahren zur Analyse von Gruppenunterschieden. Mit ihr wird versucht, eine lineare Funktion von metrisch skalierten Variablen (den Merkmalsvariablen der Elemente) zu bestimmen, mit deren Hilfe bestmöglich zwischen zwei (oder mehr) Gruppen unterschieden werden soll. Damit können im Rahmen der Kreditrisikoanalyse jene Merkmale - also Kennzahlen und Marktvariablen - bestimmt werden, die die Gruppe jener Kreditnehmer, die ihren Zahlungen nachkommen, von der Gruppe jener Kreditnehmer unterscheidet, die ihren Zahlungen nicht nachkommen können. Zusätzlich können damit neue Kreditnehmer den Gruppen zugeordnet werden, wobei die Wahrscheinlichkeit einer Fehlklassifikation minimiert wird. Daraus lässt sich auch die Wahrscheinlichkeit für die Zugehörigkeit zur Gruppe der insolventen Kreditnehmern $P_i(y_i = 1)$ ableiten, die bereits in Gleichung (1) dargestellt wurde.

Zentraler Bestandteil der Diskriminanzanalyse ist die Diskriminanzfunktion, die aus einer Linearkombination der j Merkmalsvariablen $X = (x_1, x_2, \dots, x_j)$ gebildet wird. Werden die zu schätzenden Diskriminanzkoeffizienten im Vektor β zusammengefasst, nimmt die Diskriminanzfunktion folgende allgemeine Form an:

$$(2) \quad y = X\beta.$$

Die Diskriminanzkoeffizienten sind so zu schätzen, dass die Streuung zwischen den Gruppen maximiert, die Streuung in den Gruppen hingegen minimiert wird. Als Streuung wird dabei jenes Maß definiert, das die durchschnittliche Distanz der Ausprägungen der Diskriminanzvariable y innerhalb einer Gruppe (bzw. zwischen den Gruppen) auf der Diskriminanzachse $X\beta$ misst. Ebenfalls auf der Diskriminanzachse wird der kritische Diskriminanzwert y^* bestimmt, der zur Prognostizierung der Gruppenzugehörigkeit neuer Kreditnehmer dient. Wird dieser Schwellenwert überschritten, wird der Kreditnehmer der einen, sonst der anderen Gruppe, zugewiesen.

⁶Für eine kurze Einführung vergleiche Backhaus et al. (2008) S. 181ff. und Bärbel Elpelt (1989) S. 240ff.

Backhaus et al. (2008) unterscheidet drei unterschiedliche Klassifizierungskonzepte: Das Distanzkonzept als klassischer Ansatz entspricht der Zuordnung anhand des oben erwähnten Diskriminanzwertes y^* . Formal wird jener Gruppe das Element zugeordnet, zu der die Distanz auf der Diskriminanzachse am kleinsten ist.

Das zweite Klassifizierungskonzept beruht auf Klassifizierungsfunktionen, die für jede Gruppe gesondert bestimmt werden müssen. Die Klassifizierung kann dadurch direkt auf Basis der Merkmalswerte unter Umgehung der Diskriminanzfunktion durchgeführt werden. Auch können in den gesonderten Klassifizierungsfunktionen *a priori*-Wahrscheinlichkeiten⁷ berücksichtigt werden.

Das Wahrscheinlichkeitskonzept schliesslich erlaubt zusätzlich zur Berücksichtigung der *a priori*-Wahrscheinlichkeiten die Einbeziehung der Kosten einer möglichen Fehlklassifizierung. Das ist immer dann von Bedeutung, wenn diese Kosten asymmetrisch sind, d.h. zum Beispiel wenn die Kosten der Zuordnung eines "schlechten" Kreditnehmers zur Gruppe der "guten" Kreditnehmer größer sind als jene bei Zuordnung eines "guten" Kreditnehmers zur Gruppe der "schlechten". In diesem Fall wird die Bank die zweitgenannte Fehlklassifizierung eher in Kauf nehmen. Durch Anwendung der Bayes'schen Regel werden schliesslich die Klassifizierungswahrscheinlichkeiten ermittelt.

Die Diskriminanzfunktion (2) weist eine große Ähnlichkeit zu einer linearen Regressionsfunktion auf. Tatsächlich weisen Backhaus et al. (2008) darauf hin, dass, da eine binäre Variable sich formal immer wie eine metrische behandeln lässt, eine formale Übereinstimmung zwischen Diskriminanz- und Regressionsanalyse besteht.⁸ Die modelltheoretischen Unterschiede bleiben aber wesentlich: Während im Regressionsmodell die abhängigen Variablen fix sind, sind in bei der Diskriminanzanalyse die Gruppen fix mit zufällig variierenden Merkmalen!

⁷Das sind jene Wahrscheinlichkeiten bezüglich der Gruppenzugehörigkeit, die vor Durchführung der Diskriminanzanalyse geschätzt werden können. Backhaus et al. (2008) nennen als Beispiel eine Person, von der *a priori* angenommen werden kann, eher Wähler einer großen Partei zu sein als Wähler einer Kleinpartei, einfach weil eine große Partei mehr Wähler hat!

⁸Im Zwei-Gruppen-Fall; vgl. Backhaus et al. (2008) S.199ff.

Caouette et al. (1998) beschreibt und vergleicht die wichtigsten Varianten dieses Modells: Altmans *Z-score*-Modell, das *Private Firm Z-score*-Modell, das *Nonmanufacturers' Z-score*-Modell sowie das *Zeta Credit Risk*-Modell.⁹ Exemplarisch wird die Diskriminanzfunktion von Altmans *Z-score*-Modell dargestellt. Sie nimmt folgende Form an:

$$z^d = 1,2x_1 + 1,4x_2 + 3,3x_3 + 0,6x_4 + 0,999x_5$$

mit:

- $x_1 = \textit{working capital}$ / Gesamtvermögen,
- $x_2 = \text{Kapitalrücklagen}^{10}$ / Gesamtvermögen,
- $x_3 = \text{EBIT}^{11}$ / Gesamtvermögen,
- $x_4 = \text{Eigenkapital zu Marktwerten} / \text{Langfristige Verbindlichkeiten zu Buchwerten},$
- $x_5 = \text{Umsätze} / \text{Gesamtvermögen}.$

Je höher der Wert z^d , desto geringer die Ausfallsrisikoklassifizierung des Kreditnehmers. Ein geringer oder negativer z^d Wert ist ein Indiz für die Zugehörigkeit zu einer recht hohen Ausfallsrisikoklasse.

3.3. Das lineare Wahrscheinlichkeitsmodell. Das lineare Wahrscheinlichkeitsmodell gehört, zusammen mit den weiter unten behandelten Probit- und Logit-Modellen zu den binären Response Modellen. Dabei

⁹Vgl. Caouette et al. (1998) S. 115ff.

¹⁰Einschliesslich Bilanzgewinn.

¹¹*Earnings before interest and taxes*; entspricht dem Gewinn vor Finanzergebnis und Steuern. In der Praxis wird häufig das Betriebsergebnis verwendet mit dem Argument, dass das außerordentlichem Ergebnis bei der Berechnung unberücksichtigt bleiben soll. Das ist insofern problematisch, da keine Vergleichbarkeit der Kennzahl gewährleistet ist, insbesondere da für Jahresabschlüsse nach IFRS der Ausweis eines außerordentlichen Posten in IAS 1/87 ausdrücklich verboten ist (Doralt (2009) S. 51 sowie Rohatschek und Maukner (2008) S. 18).

werden zur Schätzung der Ausfallwahrscheinlichkeit $E(y_i|\Omega_i)$ aus Gleichung (1) eine Indexfunktion, die aus der Linearkombination des Parametervektors β und dem Merkmalsvektor X_i gebildet wird, sowie der *linking*-Funktion F über die Indexfunktion, die über eine Transformation die Überleitung der linearen Funktion zu einer nicht-linearen Wahrscheinlichkeitsfunktion ermöglicht, herangezogen. Im einfachsten Fall des linearen Wahrscheinlichkeitsmodells wird auf die *linking*-Funktion verzichtet, bzw. wird $F(x) = x$ gesetzt.

Somit basiert das Modell auf einer linearen Regression und versucht mit Hilfe von Kennzahlen die Ausfallwahrscheinlichkeit zu schätzen, indem von Daten aus der Vergangenheit ein Modell generiert wird, mit dessen Hilfe auf die Rückzahlungswahrscheinlichkeit von neuen Krediten geschlossen wird.¹²

Alte Kredite (i) werden in zwei Gruppen geteilt: *defaulters* ($y_i = 1$) und *non-defaulters* ($y_i = 0$) Diese Beobachtungen werden mittels linearer Regression einer Menge von j kausalen Variablen $X_i = (x_{i1}, x_{i2}, \dots, x_{ij})$, die die quantitative Information über den Kreditnehmer widerspiegeln, zugeordnet:

$$y_i = \sum_{j=1}^n \beta_j x_{ij} + \epsilon_i,$$

Dabei ist β_j die geschätzte Bedeutung der Variable j um die Erfahrung der Rückzahlung in der Vergangenheit zu erklären, ϵ_i ist der Fehler. Die geschätzten β_j werden mit den beobachteten Merkmalsausprägungen x_{ij} eines bestimmten Kreditnehmers multipliziert. Daraus kann der erwartete Wert für y_i abgeleitet werden. Dieser Wert kann als Ausfallwahrscheinlichkeit $P(y_i = 1)$ des Kreditnehmers i interpretiert werden:

$$E(y_i|X_i) = P(y_i = 1|X_i)$$

Der Vorteil des Modells gegenüber anderen Methoden, insbesondere den Probit- und Logit-Modellen, ist die relative Einfachheit, mit der sich die Regressionsgerade bei großen Datenmengen schätzen lässt. Allerdings bemerken Davidson und MacKinnon (1993): "*In view of the much better models that*

¹²Vgl. Saunders (1997) S. 278.

are available, and the ease of estimating them using modern computer technology, this model has almost nothing to recommend it."¹³

Zwei Kritikpunkte sollen an dieser Stelle hervorgehoben werden: Zum einen wird die Interpretation der Ergebnisse problematisch, wenn y_i kleiner als Null oder größer als 1 wird, da Wahrscheinlichkeiten ermittelt werden sollen. Aufgrund der Linearität der Regressionsfunktion wird dieser Fall aber fast zwingend eintreten. Zum anderen sind nicht alle Voraussetzungen für die Schätzung der Linearregression mittels der Methode der kleinsten Quadrate erfüllt.¹⁴ Insbesondere sind aufgrund der dichotomen Eigenschaft der abhängigen Variable die Fehler ϵ_i nicht normal- sondern binomialverteilt¹⁵ und verletzen durch die von X_i abhängigen Varianz die Bedingung der Homoskedastizität.¹⁶

3.4. Das Probit-Modell. Um eine aussagekräftige Prognose zu ermöglichen, sollte die Transformationsfunktion mindestens folgende Eigenschaften erfüllen: Sie sollte nach unten mit 0 und nach oben 1 beschränkt sein, und ihrem gesamten Bereich stetig monoton wachsend sein,¹⁷ formal:

$$(3) \quad \lim_{x \rightarrow -\infty} F(x) = 0 \text{ und } \lim_{x \rightarrow \infty} F(x) = 1,$$

$$(4) \quad \frac{\partial F(x)}{\partial x} > 0.$$

Während eine lineare Funktion nicht beide Bedingungen gleichzeitig erfüllen kann, trifft genau dies für eine Großzahl von Verteilungsfunktionen zu. Diese eignen sich daher besonders als Transformationsfunktion.

Das Probit-Modell zieht die Verteilungsfunktion der Standardnormalverteilung $\Phi(x)$ heran:

$$(5) \quad \begin{aligned} P(y_i = 1) &= E(y_i | X_i) \\ &= \Phi(X_i \beta) \end{aligned}$$

¹³Davidson und MacKinnon (1993) S. 513.

¹⁴Zur linearen Regressionsanalyse siehe zum Beispiel und Bärbel Elpelt (1989) Kapitel 2.

¹⁵Vgl. Hosmer und Lemeshow (1989) S. 7.

¹⁶Vgl. Backhaus et al. (2008) S. 422.

¹⁷Vgl. Davidson und MacKinnon (1993) S. 513.

$$= \int_{-\infty}^{X_i\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt.$$

Das Probit Modell kann von einem Modell abgeleitet werden, das eine latente Variable beinhaltet.¹⁸ Eine latente Variable y_i° ist eine metrische Variable, die von den Merkmalsausprägungen X_i abhängt aber nicht direkt beobachtet werden kann, von deren Realisierung allerdings die beobachtbare dichotome Variable y_i abhängt.

Wird für die latente Variable Linearität unterstellt, kann sie wie folgt definiert werden:

$$y_i^\circ = X_i\beta + \epsilon_i,$$

mit einem standardnormalverteiltem Fehler ϵ_i :

$$\epsilon_i \sim \mathcal{N}(0, 1).$$

Die beobachtbare Variable hängt nun von der latenten Variable direkt ab:

$$y_i = \begin{cases} 1 & \text{wenn } y_i^\circ > 0, \\ 0 & \text{wenn } y_i^\circ \leq 0. \end{cases}$$

Nun entspricht die Wahrscheinlichkeit $P(y_i = 1)$ der Wahrscheinlichkeit $P(y_i^\circ > 0)$. Es gilt:

$$\begin{aligned} y_i^\circ &> 0 \\ X_i\beta + \epsilon_i &> 0 \\ \epsilon_i &> -X_i\beta. \end{aligned}$$

Da aber ϵ_i standardnormalverteilt ist, können die Wahrscheinlichkeiten wie folgt berechnet werden:

$$P(y_i = 1) = P(\epsilon_i > -X_i\beta)$$

¹⁸Vgl. Davidson und MacKinnon (1993) S. 514f und Ronning (1991) S. 8ff.

$$\begin{aligned}
 &= 1 - P(\epsilon_i \leq -X_i\beta) \\
 &= 1 - \Phi(-X_i\beta) \\
 (6) \quad &= \Phi(X_i\beta).
 \end{aligned}$$

Das Ergebnis aus Gleichung (6) entspricht genau der Spezifizierung des Probit-Modells in Gleichung (5).

Die Parameter β wird üblicher Weise mittels der *maximum likelihood*-Methode geschätzt. Dabei soll die Wahrscheinlichkeit (i.S.v. *likelihood*, nicht *probability*!), dass alle Beobachtungen gleichzeitig mit der Modellvorhersage zutreffen, maximiert werden. Genauer wird auf die *maximum likelihood*-Methode in Kapitel 5 eingegangen.

3.5. Das Logit-Modell. Das letzte hier behandelte Modell aus der Klasse der binären *response*-Modelle, und gleichzeitig jenes, das in Kapitel 5 zur Anwendung kommen wird, ist das Logit-Modell. Dabei wird als *linking*-Funktion, die die Bedingungen (3) und (4) erfüllt, die Verteilungsfunktion der logistischen Verteilung herangezogen, die, bei angenommenen Mittelwert 0 und Varianz $\sigma^2 = \frac{\pi^2}{3}$ folgende Form annimmt:

$$F(x) = \frac{1}{1 + e^{-x}}$$

Für das Logit-Modell ergibt sich somit eine Ausfallswahrscheinlichkeit $P(y_i = 1)$:¹⁹

$$\begin{aligned}
 P(y_i = 1) &= E(y_i|X_i) \\
 &= F(X_i\beta) \\
 (7) \quad &= \frac{1}{1 + e^{-X_i\beta}}
 \end{aligned}$$

Das Modell liefert ähnliche Ergebnisse wie das Probit-Modell, und kann ebenfalls mittels eines Latente-Variable-Modells hergeleitet werden. Da aber der Fehler ϵ_i im Logit-Fall logistisch verteilt ist mit Mittelwert 0 und Standardabweichung $\sigma^2 = \frac{\pi^2}{3}$, wird aufgrund der größeren Varianz die Verteilungsfunktion um den Mittelwert herum flacher verlaufen als beim

¹⁹Vgl. Davidson und MacKinnon (1993) S. 515 und Hosmer und Lemeshow (1989) S. 6.

Probit-Modell. Davidson und MacKinnon (1993) zeigen allerdings, dass bei einer Reskalierung des Logit-Modells auf eine Varianz von $\sigma^2 = 1$ die Abweichungen vom Probit-Modell kaum noch ins Gewicht fallen.

Die latente Variable nimmt eine ähnliche Form an wie im Probit-Modell:

$$y_i^\circ = X_i\beta + \epsilon_i,$$

nur dass der Fehler ϵ_i logistisch verteilt ist mit $\alpha = 0$ und $\beta = 1$. Diese Parameterausprägung der Verteilung entspricht einem Mittelwert von 0 und einer Varianz von $\sigma^2 = \frac{\pi^2}{3}$.

Die beobachtbare Variable hängt wieder von der latenten Variable direkt ab:

$$y_i = \begin{cases} 1 & \text{wenn } y_i^\circ > 0, \\ 0 & \text{wenn } y_i^\circ \leq 0. \end{cases}$$

Analog zum Probit-Modell kann nun die Wahrscheinlichkeit für $P(y_i = 1)$ berechnet und das Logit-Modell hergeleitet werden:

$$\begin{aligned} P(y_i = 1) &= P(\epsilon_i > -X_i\beta) \\ &= 1 - P(\epsilon_i \leq -X_i\beta) \\ &= 1 - \frac{1}{1 + e^{X_i\beta}} \\ (8) \qquad &= \frac{1}{1 + e^{-X_i\beta}}. \end{aligned}$$

Das Ergebnis aus Gleichung (8) entspricht der Spezifizierung des Logit-Modells aus Gleichung (7).

So wie die Probit-Funktion unterstellt also auch die logistische Regressionsfunktion einen nicht linearen Zusammenhang zwischen der Eintrittswahrscheinlichkeit der binären, abhängigen Variablen $P(y = 1)$ und den unabhängigen Variablen X_i als Modellprämisse.²⁰ Die Parameterschätzung erfolgt ebenfalls wie im Probit-Modell über die *maximum likelihood*-Methode.

²⁰Vgl. Backhaus et al. (2008), S. 423.

	Lineares Wahrscheinlichkeitsmodell	Probit-Modell	Logit-Modell
Indexfunktion	$y_i = X_i\beta_i + \epsilon_i$	$y_i = X_i\beta_i + \epsilon_i$	$y_i = X_i\beta_i + \epsilon_i$
linking-Funktion	$F(x) = x$	$F(x) = \Phi(x)$	$F(x) = \frac{1}{1+e^{-x}}$
Implizite Verteilungsannahme für ϵ_i	binomial	standardnormal	logistisch mit $\sigma^2 = \frac{\pi^2}{3}$

TABELLE 2.1. Überblick über wesentliche Modellannahmen der untersuchten binären *response*-Modelle.

Gegenüber dem Probit-Modell hat das Logit-Modell den wesentlichen Vorteil, dass die so genannten *odds*, die "Quoten", das Verhältnis der Wahrscheinlichkeit $P(y_i = 1)$ zur Gegenwahrscheinlichkeit $P(y_i = 0) = 1 - P(y_i = 1)$ unmittelbar aus dem Modell ableitbar sind. Wie noch zu zeigen ist, gilt:

$$\ln \left(\frac{P(y_i = 1)}{1 - P(y_i = 1)} \right) = X_i\beta.$$

Dieses Eigenschaft erleichtert nicht nur die Manipulation der Daten, sondern vereinfacht auch die Interpretation der Ergebnisse, wie in Kapitel 5 erläutert werden wird.

An dieser Stelle ist es sinnvoll, eine die unterschiedlichen Prämissen der drei hier untersuchten binären *response*-Modelle zusammenzufassen, wie in Tabelle 2.1 dargestellt.

3.6. Neuronale Netzwerke. Einen relativ neuen Ansatz in der Kreditrisikoklassifizierung stellen neuronale Netzwerke dar.²¹ Diese lehnen sich in der Funktionsweise an ihre biologischen Entsprechungen, den neuronalen Zellen: Diese geben einen Reiz erst weiter, wenn ein bestimmter Schwellenwert erreicht wird. Durch parallele Schaltung mehrerer Zellen können

²¹In diesem kurzen Abschnitt wird nur das MTL-Netz (*multi-layer perceptron*) vorgestellt, vgl. Backhaus et al. (2008) S. 527ff. Für eine umfangreichere Einführung in die Anwendungen neuronaler Netze in der Ökonomie siehe z.B. Zimmermann (1994), für spezifische Anwendungen zur Bewertung des Kreditrisikos siehe Kerling und Poddig (1994), Caouette et al. (1998) und Schmidt-von Rhein und Rehkugler (1994).

dadurch, auch wenn die einzelnen Neuronen nur ein binäres Signal erzeugen, nichtlineare Vorgänge abgebildet werden.

Der Aufbau eines neuronalen Netzes ist in Abbildung 2.2 dargestellt: Es besteht aus einer Eingabeschicht, mindestens einer verdeckten Schicht und einer Ausgabeschicht. Die Eingabeschicht dient nur dazu, die Eingabeparameter des Modells aufzunehmen, die Ausgabeschicht liefert das Ergebnis.

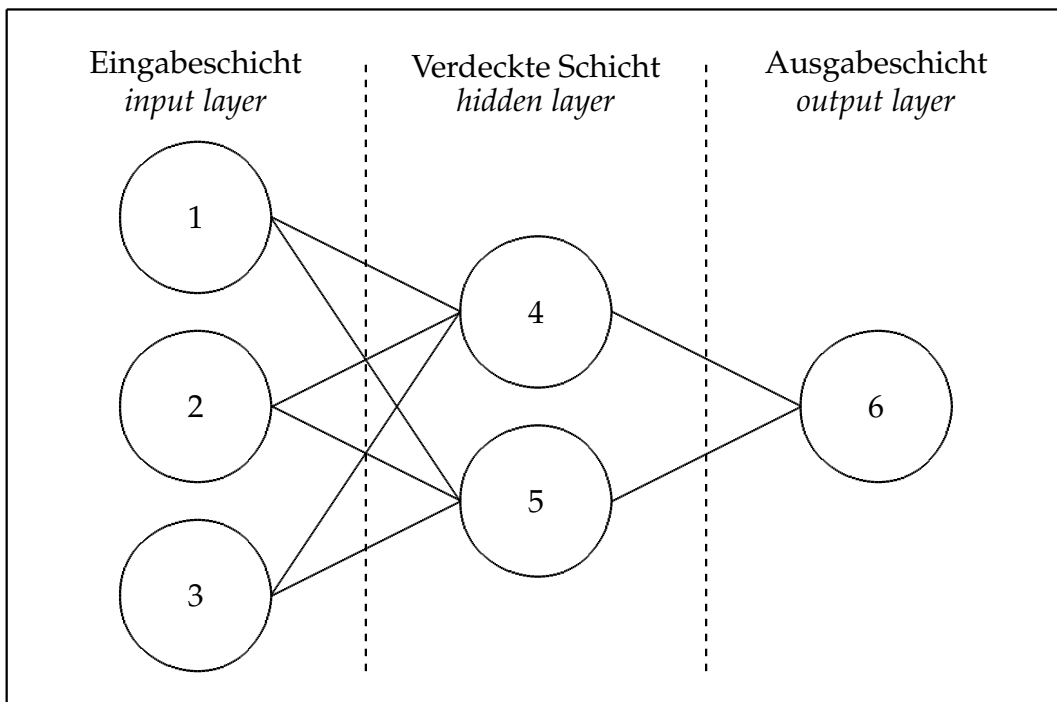


ABBILDUNG 2.2. Schematische Darstellung eines neuronalen Netzwerks. Q: Backhaus et al. (2008) S. 741

Ein Neuron der verdeckten Schicht übernimmt die Werte der vorgelagerten Schicht und bildet daraus einen gewichteten, aggregierten Wert (Nettoeingabewert). Analog zum biologischen Neuron bestimmt nun die Aktivierungsfunktion, ob und wie sehr die Zelle aktiviert wird, also ein Signal an die nachfolgende Schicht weitergibt.

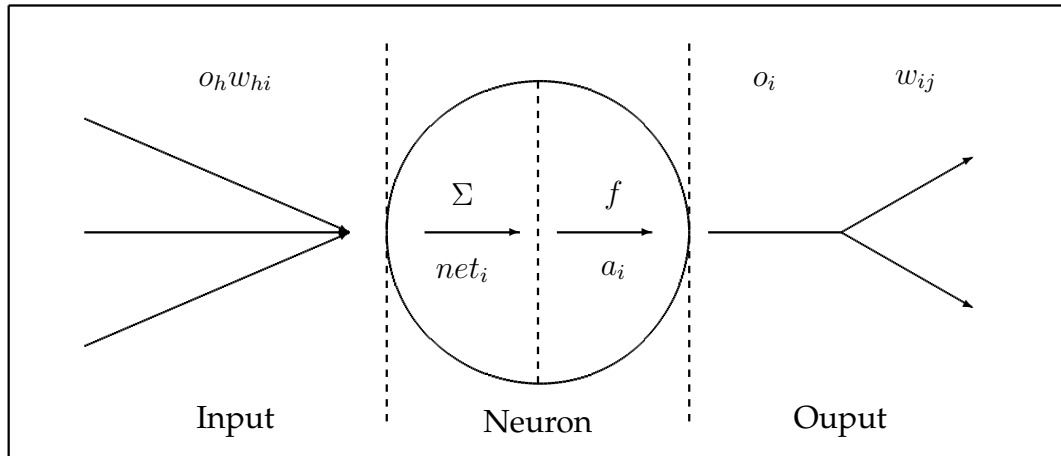


ABBILDUNG 2.3. Aufbau eines aktiven Neurons in der verdeckten Schicht. Q: Backhaus et al. (2003) S. 757

Eine allgemeine schematische Darstellung einer neuronalen Einheit ist in Abbildung 2.3 wiedergegeben: o_h steht dabei für den Output der Vorgängerzelle h , und w_{hi} für die Gewichtung des Output der Zelle h für die Bestimmung des Nettoeingabewerts net_i . In der Regel wird dabei die Summenfunktion verwendet, so dass die Propagierungsfunktion Σ (d.i. jene Funktion, mittels derer der Nettoeingabewert berechnet wird) wie folgt definiert ist:

$$net_i = \sum_{h \in H} o_h w_{hi},$$

mit H die Mengen der Indizes für die Neuronen der vorgelagerten Schicht.

Die Aktivierungsfunktion f vergleicht diesen Wert mit dem Schwellenwert a_i und bestimmt den Output o_i . Im einfachsten Fall handelt es sich um eine Schwellenwertfunktion, die wie folgt definiert wird:

$$f(net_i) = \begin{cases} 1 & \text{wenn } net_i \geq a_i, \\ 0 & \text{wenn } net_i < a_i. \end{cases}$$

Zur Spezifizierung des Modells werden, nachdem durch den Modelldesigner die Anzahl der Schichten sowie die Propagierungs- und Aktivierungsfunktionen festgelegt worden sind, mit numerischen Verfahren die Gewichte solange verändert, bis die größte Anzahl der Beobachtungen mit den

Modellvoraussagen übereinstimmen. Dieses Verfahren wird mit "lernen" bezeichnet.

Die neuronalen Netze liefern Ergebnisse, die jener der Diskriminanzanalyse sehr nahe kommen und können immer dann angewendet werden, wenn nur die Ergebnisse einer Klassifizierung erforderlich sind, da, im Gegensatz zur Diskriminanzfunktion, die eine gewisse Interpretation der Parameter erlaubt, die Neurone der verdeckten Schichten, für sich genommen, kaum eine Interpretation bezüglich des Einflusses bestimmter Eingabevariablen erlauben.²²

3.7. Hybride Systeme. Diese beruhen auf Basis von Berechnungen, Schätzungen und Simulationen. Das KMV Modell fällt zum Beispiel in diesen Bereich. Einen Vergleich von Kreditrisikomodellen aus diesem Bereich beschreiben Crouhy et al. (2000).

4. Accuracy Ratio

Alle beschriebene Modelle haben das Ziel, möglichst zuverlässig Unternehmen, denen die Insolvenz droht, von jenen zu unterscheiden, denen dieses Schicksal nicht droht. Ein wesentliches Qualitätsmerkmal eines Modells kann daher über die Trennschärfe der Ratingmethode ermittelt werden.²³

Zur Veranschaulichung der Trennschärfe kann die Kontingenztafel herangezogen werden. Bei zwei Ratingklassen (Solvenz/Insolvenz) entsteht eine 2×2 -Matrix, in der die korrekt prognostizierten Insolvenzen und die fälschlicher Weise als Insolvenz klassifizierte Solvenzen (α -Fehler oder auch *falsch-positiv* bezeichnet) in der oberen und die fälschlicher Weise als Solvenz klassifizierte Insolvenzen (β -Fehler oder auch *falsch-negativ* bezeichnet) und die korrekt prognostizierten Solvenzen in der unteren Zeile. Um die Trennschärfe des zugrunde liegenden Modells zu bestimmen, werden die "Trefferquote", d.i. das Verhältnis der richtig erkannten Insolvenzen zu allen Insolvenzen, und die "Fehlalarmquote", d.i. das Verhältnis der irrtümlich als insolvent klassifizierte Solvenzen zu allen Solvenzen, gebildet.

²²Backhaus et al. (2008), S. 738.

²³Für diesen Abschnitt, vgl. Reichling et al. (2007), S. 113ff.

Die Logitregression liefert eine Ausfallwahrscheinlichkeit, daher muss zur Klassifizierung ein *cut-off point* gewählt werden. So könnten zum Beispiel Solvenzen als jene Fälle definiert werden, für die das Modell eine Überlebenswahrscheinlichkeit von über 50% berechnet. Im Idealfall gibt es einen *cut-off point* y^* , so dass

$$\begin{aligned} P(y_i = 0) \leq y^* &\Leftrightarrow y_i = 0 \text{ und} \\ P(y_i = 0) > y^* &\Leftrightarrow y_i = 1 \end{aligned}$$

gilt. Damit ist eine perfekte Trennschärfe gegeben, die durch eine Trefferquote von 1 charakterisiert wird. Tatsächlich wird diese Trennschärfe nicht erreicht werden, da es nahe an y^* zu Überschneidungen kommt. Die Treffer- und Fehlalarmquoten hängen direkt vom gewählten *cut-off point* ab und spiegeln diese Überschneidungen wider: Wird der Wert von y^* verringert, wird die Anzahl der prognostizierten Solvenzen vergrößert und die Fehlalarmquote reduziert. Gleichzeitig sinkt aber auch die Trefferquote, weil weniger Insolvenzen richtig erkannt werden.

Werden Treffer- und Fehlalarmquoten für alle möglichen Werte y^* (im Fall der Logitregression das Intervall $[0, 1]$) berechnet, können Informationen über die Qualität der Voraussage des Modells für die gesamte Bandbreite der Risikobewertung eines bestimmten Zeitraumes gewonnen werden. Zur Veranschaulichung bieten sich 2 Modelle an: Die *receiver operating characteristic*-Kurve (ROC), sowie die *cumulative accuracy profile*-Kurve (CAP), die beide informationsökonomisch gleichwertig sind, also die Information nur unterschiedlich darstellen.²⁴

Die ROC-Kurve wurde ursprünglich entwickelt, um ein Problem der Signalübertragung zu lösen: Bei dieser entsteht in der Regel ein Rauschen. Während ein Teil des Signal eindeutig als Signal und ein Teil des Signals eindeutig als Rauschen identifiziert werden kann, gibt es dazwischen einen Bereich, wo sich Rauschen und Signal überlappen. Die Kurve wird gebildet indem für alle möglichen *cut-off points* (also alle möglichen Werte y^*) die Trefferquote auf der Abszisse gegen die Fehlalarmquote auf der Abszisse eingetragen wird. Die Kurve liegt zwischen dem Idealfall, d.i. die Trefferquote 1,

²⁴Für den formalen Beweis der Gleichwertigkeit vgl. Engelmann et al. (2003).

und der Zufallratingfunktion, bei der ein richtiger Treffer so häufig ist wie ein falscher, also die Trefferquote der Fehlalarmquote entspricht. Umso höher die Kurve liegt, umso näher liegt sie beim Optimum, und umso besser ist das zugrunde liegende Modell.²⁵

Die CAP-Kurve lehnt sich konzeptuell an der Lorenzkurve an und stellt den kumulierten Anteil insolventer Unternehmen auf der Abszisse in Relation zum kumulierten Anteil der nach Ausfallswahrscheinlichkeit geordneten Unternehmen, wie in Abbildung 2.4 dargestellt, d. h. dass für einen gegebenen Anteil x der Gesamtheit der Unternehmen die CAP-Kurve jenen Prozentsatz $y(x)$ der Zahlungsunfähigen darstellt, deren Risikobewertung gleich oder geringer ist als derer vom Anteil x . Bei einem perfekten Ratingmodell ist die CAP-Kurve zunächst ansteigend linear, da in diesem Bereich die Insolvenzen erfasst werden. Schliesslich erfasst sie 100% der Zahlungsunfähigen und ist mit 1 begrenzt (*Ideal CAP*). Da normalerweise der Anteil der Zahlungsunfähigen sehr gering ist, ist die Ideal-CAP sehr steil. Die CAP-Kurve des evaluierten Modells liegt zwischen dem CAP des perfekten Modells und der naiven Zuordnung (*random CAP*), bei der das Modell die Zuordnung in Solvenz und Insolvenz zufällig vornimmt.

Der Zusammenhang zwischen CAP und ROC wird ersichtlich, wenn eine alternative Konstruktionsmethode angewandt wird.²⁶ Zunächst sei als Prognosequote das Verhältnis der prognostizierten Insolvenzen zur Gesamtheit aller Unternehmen. Für einen bestimmten *cut-off point* gibt die Trefferquote den Anteil der richtig erkannten Insolvenzen. Da dieser Anteil für ein bestimmtes y^* notwendiger Weise auch die Insolvenzen beinhalten muss, die eine geringere Trefferquote bei einem kleineren y^* prognostiziert hätte, handelt es sich dabei tatsächlich um den kumulierten Anteil insolventer Unternehmen, die für ein gegebenes y^* prognostiziert wird. Für den gleichen *cut-off point* wird die Prognosequote berechnet. Diese gibt den Anteil der Unternehmen an, für die eine Insolvenz erwartet wird. Formal handelt es dabei um den kumulierten Anteil aller Unternehmen, deren Überlebenswahrscheinlichkeit kleiner als y^* ist, also $P(y_i = 0) < y^*$. Somit kann die CAP Kurve auch gebildet werden, indem für alle möglichen *cut-off points*

²⁵Vgl. Kleinbaum und Klein (2010) S.355 ff.

²⁶Vgl. Reichling et al. (2007) S. 119.

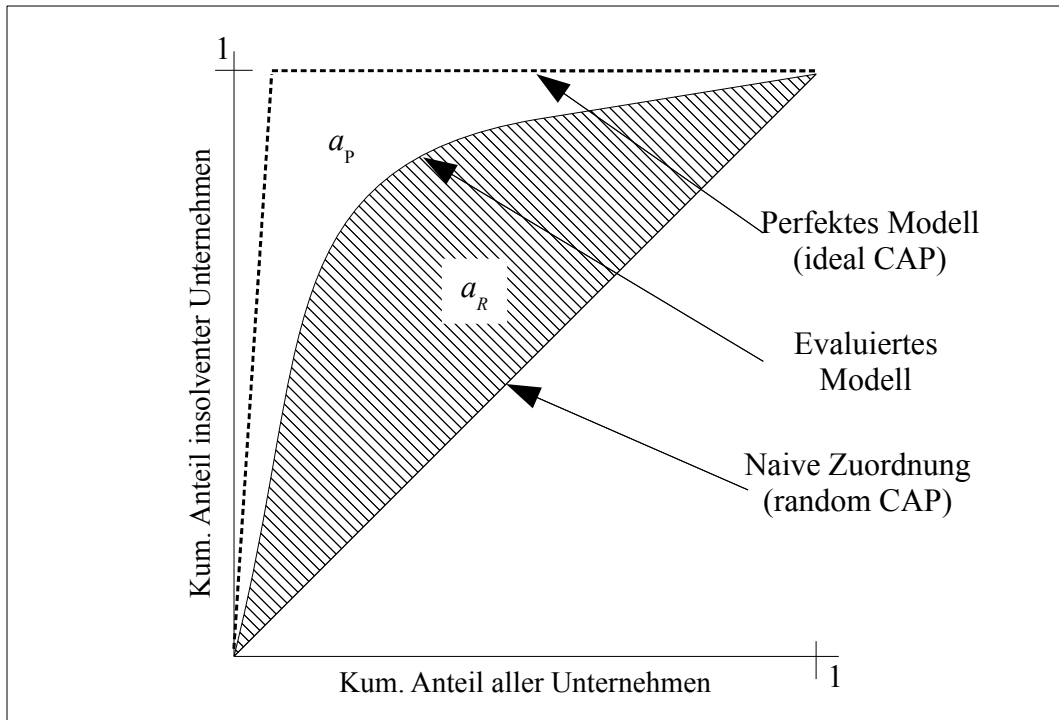


ABBILDUNG 2.4. Cumulative Accuracy Profiles. Das *accuracy ratio* ergibt sich aus dem Verhältnis der Flächen a_R und a_P .
Q: Sobehart und Keenan (2001)

(also alle möglichen Werte y^*) die Trefferquote auf der Abszisse (identisch zur ROC-Kurve) gegen die Prognosequote auf der Abszisse eingetragen wird.

Sowie sich die CAP-Kurve konzeptuell an der Lorenzkurve orientiert, gibt es analog zum GINI-Koeffizienten ein Maß für die "Nähe" einer CAP-Kurve zum *ideal CAP*. So kann die Qualität eines Ratingsystems mit dem *accuracy ratio* (AR) messen. Das *accuracy ratio* wird in Abbildung 2.4 definiert als das Verhältnis des Bereichs a_R zwischen der CAP des untersuchten Ratingmodells und der CAP der zufälligen Zuordnung und dem Bereich a_P zwischen der CAP des perfekten Modells und der CAP der zufälligen Zuordnung:²⁷

$$AR = \frac{a_R}{a_P}$$

Das *accuracy ratio* ist eine Zahl zwischen 0 und 1, wobei Werte um 1 nahezu perfekten Voraussagen entsprechen. Allerdings enthält eine einzelne Zahl

²⁷Vgl. Engelmann et al. (2003)

- aus statistischer Sicht - relativ wenig Information, sodass hauptsächlich eine vergleichende Bewertung der Trennschärfe ermittelt wird, um aus unterschiedlichen Modellen das Beste zu wählen. Weiters zeigen Engelmann et al. (2003) wie von ROC Kurven Konfidenzintervalle für *accuracy ratios* hergeleitet werden können. Wichtig ist dabei der Zusammenhang zwischen CAP und ROC, der sich der linearen Transformation des "Bereichs unter der ROC-Kurve" (AUC für *area under the curve*) um das AR zu berechnen, wieder findet.

West (1985) ermittelte mittels Faktoranalyse weit verbreitete und üblich verwendete Kennzahlen. Mit Logit Regression generierte er ein aussagekräftiges Modell zur Identifikation von Problembanken. Selbst die Ergebnisse der Faktoranalyse aus einem vorhergehenden Jahr lieferten gute Prognosen für die Bestimmung der Koeffizienten eines nachfolgenden. Daraus schließt er, dass es möglich ist ein effizientes Frühwarnsystem mittels Faktoranalyse und Logitregression zu gestalten. Auch Martin (1977) findet das Logit Modell zur Bestimmung der Ausfallswahrscheinlichkeit gut geeignet.

Platt und Platt (1991) verglichen ein Modell, das auf industriebezogenen Kennzahlen beruht mit einem Modell mit unangepassten Kennzahlen unter Verwendung der Logitregression. Die industriebezogene Spezifizierung enthält inkrementelle Information gegenüber dem Modell basierend auf unadjustierten Kennzahlen.

Für die Generierung des Kreditrisikomodells für die österreichische Hotelindustrie wird daher die Logitregression verwendet.

KAPITEL 3

Die österreichische Hotelindustrie

Dieses Kapitel befasst sich mit der Positionierung der österreichischen Hotelindustrie im weltweiten wie auch europäischen Vergleich. Des Weiteren wird auch die Struktur dieser Branche beschrieben.

Laut Statistik Austria gab es im Jahr 2011 34,6 Millionen Ankünfte in- und ausländischer Touristen und insgesamt 126 Millionen Nächtigungen in- und ausländischer Touristen.¹ Die österreichische Hotelindustrie ist nicht nur für die österreichische Wirtschaft von wesentlicher Bedeutung, sie nimmt sogar im weltweiten Vergleich einen namhaften Platz ein. So liegt Österreich mit 21,927 Millionen Tourismusankünften weltweit an zwölfter Stelle, wie aus Tabelle 3.1 ersichtlich ist.

Die "EU-Richtlinie zur Tourismusstatistik", die national durch die Tourismus-Statistik-Verordnung umgesetzt wurde, hat eine gute Vergleichsbasis für geschaffen. Der österreichische Marktanteil am europäischen Tourismus beläuft sich im Jahr 2007 laut Nächtigungen in Hotels und ähnlichen Betrieben der EU 27 auf 5,4 %. Österreich liegt damit hinter Spanien, Italien, Deutschland, Frankreich und dem Vereinigten Königreich an sechster Stelle. Eine genaue Aufschlüsselung der Nächtigungstatistiken ist in Tabelle 3.2 dargestellt: Bei den Übernachtungen pro Einwohner, an denen das Tourismusaufkommen an der Größe des Landes gemessen wird, die sogenannte Tourismusintensität, liegt Österreich mit 11,5 Übernachtungen pro Einwohner an dritter Stelle hinter Zypern (18,2) und Malta (19,4). Im Vergleich dazu liegt Österreich mit der durchschnittlichen Anzahl der Gästebetten je Betrieb mit rund 40 Betten pro Beherbergungsbetrieb vor Irland und dem Vereinigten Königreich an drittletzter Stelle vgl. Laimer und Schischeg (2009).

¹Vgl. Statistik Austria (2012).

Land	Internationale Tourismusan- künfte in 1000	Marktanteil in Prozent
Frankreich	79.300	8,6
USA	58.030	6,3
Spanien	57.311	6,5
China	53.049	5,7
Italien	42.734	4,7
Vereinigtes Königreich	30.190	3,3
Ukraine	25.449	2,8
Türkei	24.994	2,7
Deutschland	24.886	2,7
Mexiko	22.637	2,4
Malaysien	22.052	2,4
Österreich	21.927	2,4

TABELLE 3.1. Die wichtigsten Tourismusdestinationen 2008
Q: Statistik Austria (2009b)

Die Bedeutung des Tourismus in der österreichischen Volkswirtschaft zeigt die Reiseverkehrsbilanz.² Sie ist ein Teil der Leistungsbilanz, die an das Ausland erbrachte Leistungen und vom Ausland erhaltene Leistungen umfasst. Der Reiseverkehr ist der Dienstleistungsbilanz zuzuordnen. Die Entwicklung des Tourismus und die damit verbundenen Deviseneingänge korrelieren mit den Ausländerübernachtungen. Seit den siebziger Jahren stieg die Zahl der Übernachtungen von ausländischen Gästen von knapp über 70 Millionen auf rund 85,92 Millionen im Jahr 2004 an. Die Deviseneingänge stiegen nominell von rund 2 Milliarden Euro im Jahr 1979 auf über 14 Milliarden Euro im Jahr 2008 an. Der vorläufige bereinigte Eurogegenwert der Deviseneingänge aus dem Tourismus betrug im Kalenderjahr 2008 rund 14,8 Milliarden Euro. Die Ausgaben für Auslandsreisen der Österreicher lagen mit 7,7 Milliarden Euro unter den Einnahmen von Ausländern im Inland. Insofern ergaben sich Nettoeinnahmen (die Differenz aus Deviseneingängen und -ausgängen) von 7,1 Milliarden Euro. Somit trägt der positive Saldo aus dem Tourismus zu einer fast ausgeglichenen Leistungsbilanz bei.

²Vgl. Statistik Austria (2009b).

Land	Betriebe absolut	in Prozent	Betten absolut	Nächtigungen abs. in 1000	Nächtigungen pro Einwohner 2008	Betten pro Betrieb	Betten pro 1000 Einwohner
Belgien	2.013	1,0	124.811	16.196.608	1,5	62	11,7
Bulgarien	1.526	0,8	231.303	16.735.534	2,2	151,6	30,2
Dänemark	477	0,2	73.384	11.079.572	2,0	153,8	13,4
Deutschland	35.941	17,8	1.643.748	214.675.342	2,6	45,7	20
Estland	346	0,2	28.634	3.842.703	2,9	82,8	21,3
Finnland	909	0,4	119.397	15.817.079	3,0	131,3	22,6
Frankreich	18.135	9,0	1.253.962	204.268.757	3,2	69,1	19,7
Griechenland	9.207	4,5	700.933	64.085.524	5,7	76,1	62,6
Irland	4.087	2,0	156.775	28.282.000	6,5	38,4	36,0
Italien	34.058	16,8	2.142.786	254.328.742	4,3	62,9	36,1
Lettland	318	0,2	20.685	2.758.702	1,2	65,0	9,1
Litauen	348	0,2	21.871	2.591.262	0,8	62,8	6,5
Luxemburg	273	0,1	14.559	1.437.539	3,0	53,3	30,3
Malta	160	0,1	39.985	7.916.648	19,4	249,9	97,8
Niederlande	3.196	1,6	200.254	34.158.500	2,1	62,7	12,2
Österreich	14.204	7,0	573.726	79.166.924	9,5	40,4	69,0
Polen	2.443	1,2	190.387	24.306.848	0,6	77,9	5,0
Portugal	2.031	1,0	264.747	39.76.583	3,7	130,4	25,0
Republik Zypern	735	0,4	87.804	14.298.478	18,2	119,5	112,0
Rumänien	4.163	2,1	228.123	19.755.878	0,9	54,8	10,6
Schweden	1.893	0,9	207.439	25.416.188	2,8	109,6	22,7
Slowakei	1.249	0,6	67.178	7.233.054	1,3	53,8	12,4
Slowenien	396	0,2	33.040	5.545.819	2,7	83,4	16,4
Spanien	17.827	8,8	1.642.417	271.689.481	6,1	92,1	36,6
Tschechien	4.559	2,3	248.077	27.043.558	2,6	54,4	24,0
Ungarn	1.999	1,0	154.088	16.297.146	1,6	77,1	15,3
Ver. Königreich	39.860	19,7	1.245.064	169.483.745	2,8	31,2	20,4
EU-27	202.353	100	11.715.177	1.578.148.214	3,2	57,9	23,6

TABELLE 3.2. Tourismus in der europäischen Union, Q: Laimer und Schischeg (2009) S. 74.

Die österreichische Hotelindustrie besteht in erster Linie aus Klein- und Mittelbetrieben. Vor allem in ländlichen Gebieten sind die Hotels Familienbetriebe.³ Im Durchschnitt betrug die Bettenzahl je Betrieb rund 15 Betten. Nach Unterkunftsarten ergibt sich jedoch ein anderes Bild: In Hotels und ähnlichen Betrieben betrug die Bettenanzahl rund 40 Betten, in Privatquartieren und privaten Ferienwohnungen/-häusern hingegen nur 7 Betten je Betrieb.

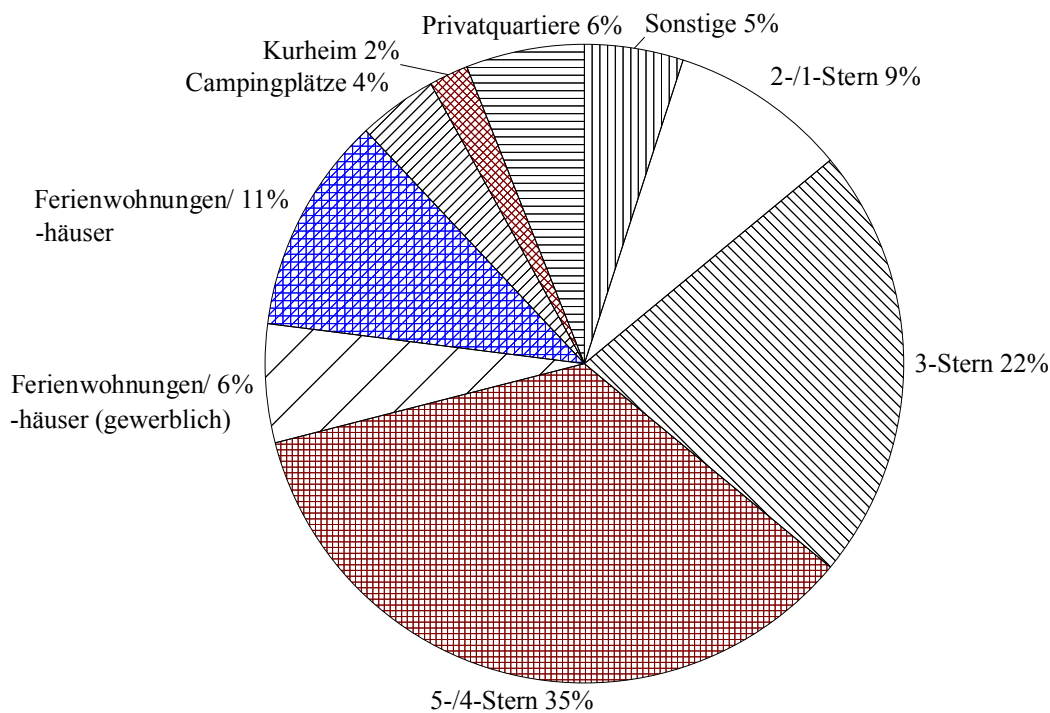


ABBILDUNG 3.1. Übernachtungen nach Unterkunftsarten im Kalenderjahr 2008. Q: Laimer und Schischeg (2009) S. 58.

Wie aus Graphik 3.1 ersichtlich, entfallen rund 50 Prozent aller Betten auf Hotelbetriebe und gewerbliche Ferienwohnungen/-häuser. Ein knappes Drittel dagegen wird von privat geführten Betrieben bereitgestellt, 22% entfallen auf sonstige Beherbergungsbetriebe (z. B. Jugendherbergen und Jugendgästehäuser, bewirtschaftete Schutzhütten, Campingplätze, Kinder-

³Vgl. Stomper und Zulehner (2004).

und Jugenderholungsheime. In Österreich ist die Bettenanzahl in Privatquartieren rückläufig, der Trend in der 5-/4-Stern Kategorie jedoch kontinuierlich steigend.

Hotels und ähnliche Betriebe sind mit 65% Übernachtungsanteil das wichtigste Segment der Unterkunftsarten, wobei 5-/4-Stern Betriebe einen Anteil von 35% einnehmen, gefolgt von den 3-Stern-Betrieben mit 22%. Die zweitwichtigste Unterkunftsart nach Übernachtungen sind private Ferienwohnungen/-häuser mit einem Anteil von 11%.⁴ Im Kalenderjahr 2008 wurden insgesamt 126,7 Millionen Nächtigungen erzielt, das sind 4,3% mehr im Vergleich zu 2007.

Die Betriebe werden meist direkt von den Eigentümern geführt. Im Falle von Zahlungsunfähigkeit werden die Hotels selten sofort geschlossen, da sich für das Anlagevermögen der Hotels nur schwierig profitable Alternativen finden.

⁴Vgl. Laimer und Schischeg (2009) S. 57.

KAPITEL 4

Kennzahlendefinition und Datenerhebung

In diesem Kapitel werden zunächst die für die weitere Analyse relevanten Kennzahlen definiert, im zweiten Teil werden die erhobenen Datensätze beschrieben

1. Definition der Kennzahlen

Wie aus dem vorangegangenen Abschnitt ersichtlich, sind Kennzahlen die Basis sämtlicher Kreditrisikomodelle. Daher sollen zunächst wichtige Kennzahlen definiert und klassifiziert werden. In weiterer Folge ist die zentrale Frage der *financial ratio analysis* eine möglichst effiziente Auswahl an Kennzahlen zu treffen, um die Aktivitäten des Unternehmens bestmöglich abzudecken.

1.1. Überblick zur Klassifizierung von Kennzahlen. Chen und Shimerda (1981) weisen auf über hundert verschiedene Kennzahlen hin, von denen alleine 65 Finanzkennzahlen (*accounting ratios*) sind. Die Einteilungen der Kennzahlen variieren von Studie zu Studie. Grundsätzlich unterscheidet man folgende Ansätze in der Klassifizierung der Kennzahlen: den pragmatischen Empirismus, den deduktiven Ansatz, den induktiven Ansatz sowie den "confirmatory approach".¹

Der pragmatische Empirismus basiert auf einer subjektiven Klassifizierung der Kennzahlen aufgrund der Erfahrung bzw. Ansichten der Autoren. Verschiedene Autoren verwenden unterschiedliche Kategorien. Zu den grundlegenden Einteilungen gehören Rentabilität (*profitability*), Kapitalstruktur (*capital structure*) und Liquidität (*liquidity*).

Der deduktive Ansatz geht auf das Du Pont Schema aus dem Jahre 1919 zurück, bei dem der *ROI* (*return on investment*) mittels Kennzahlenanalyse in seine Bestandteile zerlegt und weiter entwickelt wird:

¹Vgl. Salmi und Martikainen (1994).

$$ROI = \frac{E}{V} = \frac{\frac{E}{U}}{\frac{U}{V}} = \dots,$$

mit E das Ergebnis,² V das Vermögen und U der Umsatz des Unternehmens. Die einzelnen Komponenten können so lange weiterentwickelt werden, bis rechts nur mehr Größen stehen, die direkt aus der Kostenrechnung bzw. der Buchhaltung entnommen werden (z.B. Umsatz gleich Menge mal Preis). Dieser Ansatz wurde mit dem *confirmatory approach* verbunden, da die induktiven Studien keine konsistente Klassifizierung von Kennzahlen mit sich gebracht haben. Infolgedessen wurde bei späteren Studien eine a priori Klassifizierung angenommen und dann diese empirisch getestet.

Der induktive Ansatz legt den Schwerpunkt auf statistische Methoden. Empirische Ergebnisse sind Grundlage für die Klassifizierung der Kennzahlen. Die Faktoranalyse wird häufig verwendet um eine große Anzahl von Kennzahlen in Kategorien einzuteilen, die möglichst umfassend die betrieblichen Tätigkeiten abdecken. Die Stabilität der Ergebnisse über mehrere Studien und über mehrere Jahre innerhalb einer Studie ist zentrales Thema des induktiven Ansatzes.

Chen und Shimerda (1981) untersuchen, welche der vielen Kennzahlen analysiert werden sollen. Eine Menge der Kennzahlen korrelieren miteinander. Einige werden voneinander abgeleitet. Eliminiert man diese Überschneidungen, kann man eine sinnvolle Zusammenstellung finden. Inkludiert man mehr als eine Kennzahl pro Kategorie, kann es zur Multikollinearität zwischen den Kennzahlen kommen, und diese stört das Verhältnis zwischen abhängiger und unabhängigen Variablen.

Die Darstellung der Gruppen der Kennzahlen erfolgte nach Hayden (2002). Sie wurden in die zehn gebräuchlichsten und meistzitierten Kreditrisikofaktorengruppen *leverage* (Finanzierungsanalyse), *debt coverage*, *liquidity* (Liquidität), *activity* (Analyse des *working capital*), *productivity*, *turnover* (Umschlagshäufigkeiten), *profitability* (erfolgswirtschaftliche Analyse), *firm size*

²In der Regel wird das Ergebnis vor Zinsen aber nach Steuern herangezogen. Soll allerdings der steuerliche Aspekt außer acht gelassen werden (z.B. bei der Analyse unterschiedlicher Teilbereiche eines Konzerns), wird der *EBIT* verwendet.

(Unternehmensgröße), *growth rates* (Wachstumsraten) und *leverage development* (Veränderungen der Kapitalstruktur) eingeteilt.³

1.2. Leverage. Bilanzanalytische Kennzahlen, die als Relationen zwischen verschiedenen Posten der Passivseite oder verschiedenen Posten der Aktivseite definiert sind, werden als vertikale Bilanzkennzahlen bezeichnet. Die in finanzwirtschaftlicher Sicht wichtigste vertikale Bilanzkennzahl ist der Verschuldungsgrad.⁴

Der Verschuldungsgrad misst die Fähigkeit des Unternehmens, unvorhergesehenen Ereignissen standzuhalten. Für die Definition dieser Kennzahl gibt es zwei Varianten, VS_1 , auch *gearing*, und VS_2 , auch Anspannungsgrad bezeichnet:

$$VS_1 = \frac{FK}{EK}$$

$$VS_2 = \frac{FK}{V} = \frac{FK}{EK + FK}$$

mit FK als Bezeichner für das Fremdkapital und EK für Eigenkapital. In beiden Fällen wird der gleiche Sachverhalt gemessen:

$$VS_1 = \frac{VS_2}{1 - VS_2}.$$

Die Verschuldung kann auch durch weitere Maßeinheiten erfasst werden: durch die Eigenkapitalquote (*equity ratio*) und *liabilities structure ratios*.

Der Verschuldungsgrad als Quotient aus zwei Bilanzgrößen definiert, ist abhängig von der in der Bilanz vorgenommenen Bewertung.

Die Kennzahlen, die den Anteil der Schulden im Verhältnis zum Vermögen messen, korrelieren positiv mit dem Ausfallrisiko, Kennzahlen, die den Eigenkapitalanteil messen negativ.

Folgende Kennzahlen aus dieser Gruppe werden für die spätere Analyse ermittelt:

³Um Verwechslungen zu vermeiden, werden für die Kategorisierung die englischen Begriffe von Hayden (2002) beibehalten.

⁴Vgl. Franke und Hax (2004).

$$\begin{aligned}
 k_1 &= \frac{FK}{V} \\
 k_2 &= \frac{EK}{V} \\
 k_3 &= \frac{FK_L}{V} \\
 k_4 &= \frac{FK_B}{V} \\
 k_5 &= \frac{FK_B}{FK}
 \end{aligned}$$

Die Indices L und B stehen für "langfristig" und "Bank" respektive: FK_L entspricht den langfristigen Verbindlichkeiten, FK_B den Bankverbindlichkeiten.

1.3. Debt Coverage. Das Verhältnis zwischen $EBIT$ und Zinsaufwand z misst die *Debt coverage*. Unternehmen, die nicht genügend Gewinn aus ihren Aktivitäten erzielen können, um die Zinsen zahlen zu können, haben ein erhöhtes Ausfallrisiko. Die *debt coverage* berechnet sich wie folgt:

$$k_6 = \frac{EBIT}{z}$$

1.4. Liquidity. Es handelt sich dabei um eine sehr gebräuchliche Kennzahlengruppe und kann durch viele Kennzahlen gemessen werden. Mit diesen Bilanzkennzahlen versucht man, sich ein Urteil zu bilden, in welchem Umfang die kurzfristigen Verbindlichkeiten am Bilanzstichtag gedeckt sind, einerseits durch Zahlungsmittel, und darüber hinaus auch durch andere Vermögensgegenstände, bei denen in Kürze mit Freisetzung des darin gebundenen Kapitals gerechnet werden kann. Die drei fundamentalen Liquiditätskennzahlen sind die Liquiditäten 1., 2., und 3. Grades (LQ_1 , LQ_2 und LQ_3), die auch unter den englischen Begriffen *cash ratio*, *quick ratio* und *current ratio* gebräuchlich sind (C steht hier für Zahlungsmittel und -äquivalente, R für Debitorenforderungen und der Index K für "kurzfristig"):

$$LQ_1 = \frac{C}{FK_K}$$

$$LQ_2 = \frac{C + R}{FK_K}$$

$$LQ_3 = \frac{V_K}{FK_K}$$

Bei genauerer Betrachtung der Kennzahlen sieht man, dass durchaus nicht alle vom Unternehmen zu leistenden Zahlungen in den kurzfristigen Verbindlichkeiten enthalten sind, hinzu kommen Zahlungen, die zur Aufrechterhaltung des Betriebes unumgänglich sind.

Generell gilt die Hypothese, je höher die Liquidität, z. B. Barmittel oder andere liquide Positionen, bzw. je geringer die kurzfristigen Verbindlichkeiten, desto geringer ist die Ausfallswahrscheinlichkeit. Bei den Liquiditätskennzahlen, die sich auf die Umsatzerlöse beziehen, kann man noch einen anderen Effekt feststellen: Je höher der Umsatz, desto geringer die Ausfallswahrscheinlichkeit, d.h. je kleiner der Kehrwert des Umsatzes, desto geringer das Kreditrisiko.

Folgende Kennzahlen aus dieser Gruppe werden für die spätere Analyse ermittelt (*vor* steht für Vorräte):

$$k_7 = \frac{V_K}{FK_K}$$

$$k_8 = \frac{V_K}{FK}$$

$$k_9 = \frac{V_K - FK_K}{V}$$

$$k_{10} = \frac{FK_K}{V}$$

$$k_{11} = \frac{V_K}{V}$$

$$k_{12} = \frac{V_K - FK_K}{U}$$

$$k_{13} = \frac{V_K}{U}$$

$$k_{14} = \frac{C + R}{U}$$

$$k_{15} = \frac{V_K - FK_K}{FK_K}$$

$$k_{16} = \frac{V_K - vor}{FK_K}$$

1.5. Activity Ratios. Das sind Kennzahlen, die Aspekte des Unternehmens beschreiben, die keine direkte Verbindung zum Ausfallrisiko haben, aber dennoch wichtige Informationen beisteuern. Die hier verwendeten Kennzahlen zeigen u.a. die Fähigkeit der Kunden auf, ihren finanziellen Verpflichtungen nachzukommen, z. B. anhand der Forderungen aus Lieferungen und Leistungen im Verhältnis zu den Umsatzerlösen bzw. werden die Zahlungsgewohnheiten des Unternehmens anhand der kurzfristigen Verbindlichkeiten evaluiert. Ein Unternehmen mit Liquiditätsproblemen hat höhere kurzfristige Verbindlichkeiten. Die Ausfallswahrscheinlichkeit steigt mit diesen Kennzahlen. Ausnahme ist allerdings die Kennzahl "Debitorenforderungen/Fremdkapital". In diesem Fall heisst eine höhere Kennzahl, dass ein höherer Anteil der unternehmenseigenen Schulden durch Forderungen gedeckt werden kann. *Activity ratios*, in deren Nenner Vorräte vorkommen, korrelieren positiv mit der Ausfallswahrscheinlichkeit, da wachsende Vorräte sowohl höhere Lagerkosten mit sich bringen als auch die Liquidität einschränken. Für diese Arbeit relevante Kennzahlen sind (*ma* definiert den Materialaufwand):

$$k_{17} = \frac{vor}{U}$$

$$k_{18} = \frac{vor}{ma}$$

$$k_{19} = \frac{R}{U}$$

$$k_{20} = \frac{R}{FK}$$

$$k_{21} = \frac{R}{ma}$$

$$k_{22} = \frac{R}{vor}$$

1.6. Productivity. Die zwei wesentlichen Faktoren dieser Kennzahlengruppe sind Personal- und Materialkosten im Vergleich zum Umsatz. Je höher diese Kosten sind, desto höher das Ausfallrisiko. Bezeichnet *per* die

Personalkosten, können die relevanten Kennzahlen aus dieser Gruppe wie folgt definiert werden:

$$k_{23} = \frac{per}{U}$$

$$k_{24} = \frac{U - ma}{per}$$

1.7. Turnover. Vermögensumschlagshäufigkeit gibt an, mit welcher Effizienz das verfügbare Kapital verwendet wird. Je höher der Wert dieser Kennzahl, desto eher ist es möglich, hohe Gewinne mit relativ geringen Investitionen zu erzielen und hat somit einen positiven Einfluß auf die Liquidität des Unternehmens und verringert somit das Ausfallsrisiko. Die Vermögensumschlagshäufigkeit wird wie folgt definiert:

$$k_{25} = \frac{U}{V}$$

1.8. Profitability. Die erfolgswirtschaftliche Analyse wird mit Kennzahlen durchgeführt, die entweder das Ergebnis relativ zum Anlagevermögen (Gesamtkapitalrentabilität) oder zu den Umsätzen (Umsatzrentabilität) darstellen. Je höher die Profitabilität desto desto länger dauert es, bis Verluste durch erhöhte Kosten oder durch sinkende Erlöse entstehen. Die untersuchten Kennzahlen aus dieser Gruppe sind:

$$k_{26} = \frac{EBIT}{V}$$

$$k_{27} = \frac{EBIT}{U}$$

1.9. Size. Umsätze und Gesamtkapital reflektieren die Größe eines Unternehmens. Beide Positionen werden durch den Verbraucherpreisindex (p^*) dividiert, um diese der Inflation anzupassen. Daraus ergeben sich folgenden Kennzahlen:

$$k_{28} = \frac{V}{p^*}$$

$$k_{29} = \frac{U}{p^*}$$

1.10. Growth Rates. Das Verhältnis zwischen der Wachstumsrate und der Ausfallsrate ist nicht eindeutig. Generell ist es natürlich besser für Unternehmen zu wachsen als zu schrumpfen. Andererseits kann es durch großes Wachstum zu großen Herausforderungen für das Management kommen. Untersucht wird das Umsatzwachstum (das Superskript t definiert das laufende, $t - 1$ das vorangegangene Jahr):

$$k_{30} = \frac{U^t}{U^{t-1}}$$

1.11. Change of Leverage. Während die Kennzahlen aus der Gruppe *leverage* einen punktuellen Zustand widerspiegeln, kann es sinnvoll sein, den Trend über mehrere Jahre zu betrachten. Es ist zu untersuchen ob eine Erhöhung des Verschuldungsgrades zu einer höheren Ausfallswahrscheinlichkeit führt. Für die Kennzahlen k_1 und k_4 werden diese Trends gebildet:

$$k_{31} = \frac{FK^t/V^t}{FK^{t-1}/V^{t-1}}$$

$$k_{32} = \frac{FK_B^t/V^t}{FK_B^{t-1}/V^{t-1}}$$

2. Daten und beschreibende Statistik

Die Daten für die Analyse sind von drei österreichischen Banken, der Nationalbank, der KMU Forschung Austria sowie der ÖHT zur Verfügung gestellt worden zur Verfügung gestellt worden. Der originäre Datensatz besteht aus 654 Hotels. Von diesen sind 74 als Ausfall klassifiziert. Als Ausfall gelten jene Unternehmen, die von den Banken der Gruppe der "Restrukturierer" zugeordnet wurden, z. B. bei Schuldenerlass, Zahlungsaufschub von Darlehen, Zinsen oder Gebühren, oder wenn diese in Konkurs gegangen sind.⁵

⁵Die formale Definition eines Ausfalls für das vorliegende Modell wird in Kapitel 5 gegeben.

Anzahl der Beobachtungen	Anzahl der Unternehmen
1	228
2	184
3	94
4	96
5	23
6	18
7	7
8	4

TABELLE 4.1. Übersicht der Beobachtungen. Q: Hayden (2002)

Die Beobachtungen betreffen die Jahre 1992 bis 2001, wobei der Grossteil aus den letzten drei Jahren stammt. Tabelle 4.1 erfasst die Anzahl der Beobachtungen.

Die unterschiedliche Beobachtungsdauer ergibt sich aus der Herkunft der zugrundeliegenden Datensätze. Die jeweiligen Informationen über ein Hotel sind in einer Gruppe zusammengefasst. Nur Hotels mit mindestens zwei Beobachtungen sind zur Analyse verwendet worden. Außerdem sind nur Hotels mit vollständigen Bilanzen verwendet worden. Vorläufige Bilanzen, sowie ein Beobachtungszeitraum von weniger als einem Jahr wurden nicht berücksichtigt.

414 Hotels entsprechen diesen Kriterien. Durchschnittlich gibt es 3,2 Beobachtungen pro Gruppe. Insgesamt besteht der Datensatz aus 1311 Hoteljahren.

Die Daten sind Paneldatasets. Sie enthalten Informationen über unterschiedliche Unternehmen zu unterschiedlichen Zeitpunkten. Genau genommen handelt es sich um inkomplette Panel Data sets, da nicht alle Unternehmen den gesamten Beobachtungszeitraum abdecken.

Tabelle 4.2 zeigt die Mittelwerte und Standardabweichungen der Kennzahlen. Damit lassen sich bereits Aussagen bezüglich branchenspezifischer Eigenschaften der Betriebe in der Tourismusindustrie: Zum einen liefert die Eigenkapitalquote (Kennzahl k_2) einen negativen Mittelwert, was auf eine massive Überschuldung der Betriebe deutet. Auffallend ist auch, dass die *debt coverage ratio* (Kennzahl k_6) sehr klein ist.

Kennzahl	Definition	Beobachtungen	Mittelwert	σ^2	Min	Max
k_1	FK/V	1541	.9748972	.3133539	4012683	1.800954
k_2	EK/V	1541	-.1075652	.4027505	-.9029673	.5995141
k_3	FK_L/V	1541	.615276	.4278714	0	1.537642
k_4	FK_B/V	1541	.6342939	.403236	0	1.516023
k_5	FK_B/FK	1539	.6143792	.3075201	0	.956305
k_6	$EBIT/z$	1539	.1203923	.1405037	-.1159091	.5178987
k_7	V_K/FK_K	1533	6.273008	12.85061	.0579719	50.81824
k_8	V_K/FK	1539	.1825565	.2740695	.0019672	1.031579
k_9	$(V_K - FK_K)/V$	1541	-.2039681	.252225	-.7927442	.1899077
k_{10}	FK_K/V	1541	.3442017	.2941251	.0399478	1
k_{11}	V_K/V	1541	.1364059	.1811075	.0022816	.6604958
k_{12}	$(V_K - FK_K)/U$	1539	-.4207696	.6379117	-2.474905	.2975824
k_{13}	V_K/U	1539	.2216239	.3167438	.0031038	1.282051
k_{14}	$(C + R)/U$	1503	.1900402	.2973954	-.0309111	1.158154
k_{15}	$(V_K - FK_K)/FK_K$	1539	-.5184937	.548564	-.9841401	1.00944
k_{16}	$(V_K - vor)/FK_K$	1503	.0841186	.7785376	-.98178	1.824035
k_{17}	vor/U	1503	.0230998	.0199614	0	.0764578
k_{18}	vor/ma	1423	.1528717	.1230385	.0064464	.4760533
k_{19}	R/U	1539	.0570396	.0641999	0	.2496393
k_{20}	R/FK	1539	.0475976	.0687245	0	2622127
k_{21}	R/ma	1459	.4309398	.5187655	0	1.926027

k_{22}	R/vor	1368	3.823097	5.050127	0	19.45639
k_{23}	per/U	1539	.3331651	.1207524	.0729865	.566765
k_{24}	$(U - ma)/per$	1505	2.712663	1.146593	1.351413	6.075534
k_{25}	U/V	1541	.7718493	.6596456	.0701583	2.593229
k_{26}	$EBIT/V$	1541	.029918	.073134	-.1357485	.1812927
k_{27}	$EBIT/U$	1539	.0400681	.1381288	-.3347237	.2545572
k_{28}	V/p^*	1566	502224.7	944485.6	951.7134	3512806
k_{29}	U/p^*	1566	314554.4	530572.8	777.6866	1917169
k_{30}	U^t/U^{t-1}	884	1.049642	.1684815	.7867514	1.523907
k_{31}	$(FK^t/V^t)/(FK^{t-1}/V^{t-1})$	887	1.013917	.101278	.8237545	1.307291
k_{32}	$(FK_B^t/v^t)/(FK_B^{t-1}/V^{t-1})$	810	1.026497	.2916375	.4169015	1.826241

TABELLE 4.2. Mittelwerte und Standardabweichungen der erhobenen Kennzahlen. Q: Hayden (2002)

Ein weiteres Merkmal der Tourismusbranche ist ein negatives *working capital* ($V_K - FK_K$), d.h., dass es den Unternehmen gelingt, sich über kurzfristige Verbindlichkeiten (z.B. aus Lieferungen und Leistung) zu finanzieren. Das liegt daran, dass das Umlaufvermögen gering gehalten werden kann, da wenig Vorräte erforderlich sind und auch Forderungen gegenüber den Gästen in der Regel sofort beglichen werden.

Eine letzte interessante Eigenschaft der Hotelindustrie ist ihre personalintensive Struktur, die leicht aus den Kennzahlen k_{23} und k_{24} gelesen werden kann: Da die Umsatzrentabilität nur 4% im Mittel beträgt (Kennzahl k_{27}) ist leicht ersichtlich, dass bei per/U von einem Drittel, die Personalaufwendungen mindestens ein Drittel der Gesamtaufwendungen ausmachen.

Es sind diese Charakteristika der Branche, hoher Verschuldungsgrad, geringe Rentabilität und hohe Personalintensität, die ein eigenes, zuverlässiges Kreditrisikomodell erforderlich machen.

KAPITEL 5

Analyse

Das Kreditrisikomodell für den österreichischen Hotelsektor wird mithilfe einer Logitregression auf Basis von Kennzahlen entwickelt.

1. Ausfall

Um mit Hilfe des Logit-Modells und der bereits definierten Kennzahlen die Ausfallswahrscheinlichkeit bestimmen zu können, wird die Definition des Ausfalls des Basel Committee on Banking Supervision Basel Committee on Banking Supervision (2001) übernommen. Demnach tritt ein Ausfall ein, wenn

- der Schuldner seinen Verpflichtungen nicht zur Gänze nachkommt (Kapitaltilgung, Zinsen oder Gebühren)
- ein Kreditverlustfall in Bezug auf eine Verpflichtung des Schuldners eintritt, wie z. B. *charge off*, Restrukturierung inkl. Aufschub von Kapitaltilgung, Zinsen oder Gebühren
- der Schuldner mit einer Kreditverpflichtung 90 Tage überfällig ist
- der Schuldner Konkurs angemeldet hat oder ähnliche Absicherung vor Gläubigern getroffen hat.

2. Selektion der Kennzahlen

Die Auswahl der Kennzahlen erfolgte nach Hayden (2002). Die Kennzahlen, die mithilfe des Datensatzes berechnet werden konnten, wurden in die zehn gebräuchlichsten und meistzitierten Kreditrisikofaktorengruppen *Leverage*, *Debt Coverage*, *Liquidity*, *Activity*, *Productivity*, *Turnover*, *Profitability*, *Firm Size*, *Growth Rates* und *Leverage Development* eingeteilt.

In Tabelle 5.1 sind alle Kennzahlen, geordnet nach den jeweiligen Gruppen dargestellt. In der vierten Spalte ist die Abhängigkeit zwischen Kennzahl und Ausfallswahrscheinlichkeit dargestellt. Ein + bedeutet eine Erhöhung

der Kennzahl führt zu einer Erhöhung des Ausfallsrisikos, ein - bedeutet trotz Erhöhung der Kennzahl eine Verringerung des Ausfallsrisikos.

Für eine effiziente Spezifikation des multivariaten Logitregressionsmodell müssen nach Auswahl der Kennzahlen im nächsten Schritt die aussagekräftigsten Kennzahlen pro Kreditrisikofaktor herausgefunden werden.

Dies geschieht im Rahmen der Maximum Likelihood Schätzung. Grob gesagt sucht man mit dieser Schätzung jene Werte für die unbekannt Parameter b_j , die mit der größten Wahrscheinlichkeit den bereits beobachteten Datensatz erhalten lassen. Die Maximum Likelihood Schätzung erfolgt mit Stata 7.0.

Die Likelihood-Funktion wird i. d. R. mit Hilfe des Newton-Raphson-Algorithmus maximiert (Vgl. Backhaus et al. (2008)): Schritt 1: Es werden zunächst Schätzwerte für die Logit-Koeffizienten angenommen, wobei diese Startwerte über eine Kleinste Quadrate Schätzung ermittelt werden können. Schritt 2: Für einen beliebigen Beobachtungsfall i wird mit der aus Schritt 1 gewonnenen Logit-Koeffizienten der Logit berechnet und die Wahrscheinlichkeit bestimmt. Schritt 3: Für diesen gewählten Fall wird der LogLikelihood-Wert berechnet. Schritt 4: Die Schritte 2 und 3 werden für alle Beobachtungsfälle durchgeführt um eine Gesamt-LogLikelihood-Funktion zu bestimmen Schritt 5: Schritte 2 bis 4 werden mit anderen Werten von β_j wiederholt. Schritt 6: Die Gesamt-LogLikelihood Funktionen werden miteinander verglichen und die Regressionskoeffizienten so lange verändert, bis keine deutliche Steigerung der Gesamt-LogLikelihood-Funktion mehr möglich ist.

1

Die Interpretation der Regressionskoeffizienten gestaltet sich nun als schwierig, da kein linearer Zusammenhang zwischen den unabhängigen Variablen X_i und den über die logistische Funktion bestimmten Wahrscheinlichkeiten $P_i(y_i = 1)$ besteht. Eine lineare Verknüpfung gilt lediglich für den y_i^0 -Wert, der aggregierten Einflussstärke.² Änderungen in den

¹Eine detaillierte Einführung in *Maximum Likelihood* Techniken bietet z. B. Kleinbaum und Klein (2010) S 106 ff.

²Vgl. Backhaus et al. (2008) S. 256 ff.

Kz.	Definition	Kreditrisikoklasse	Hyp.	Litteratur
k_1	FK/V	Leverage	+	a, c, d, e, f
k_2	EK/V	Leverage	-	a, c
k_3	FK_L/V	Leverage	+	d, e
k_4	FK_B/V	Leverage	+	a, c
k_5	FK_B/FK	Leverage	+	d
k_6	$EBIT/z$	Debt Coverage	-	a, f
k_7	V_K/FK_K	Liquidity	-	a, c, d, e, f
k_8	V_K/FK	Liquidity	-	d
k_9	$(V_K - FK_K)/V$	Liquidity		
k_{10}	FK_K/V	Liquidity	-	d, e
k_{11}	V_K/V	Liquidity	-	d, e
k_{12}	$(V_K - FK_K)/U$	Liquidity	-/+	
k_{13}	V_K/U	Liquidity	-/+	
k_{14}	$(C + R)/U$	Liquidity	-/+	
k_{15}	$(V_K - FK_K)/FK_K$	Liquidity	-	d
k_{16}	$(V_K - vor)/FK_K$	Liquidity		
k_{17}	vor/U	Activity	+	d, e
k_{18}	vor/ma	Activity	+	a, c
k_{19}	R/U	Activity	+	a, e
k_{20}	R/FK	Activity	-	d
k_{21}	R/ma	Activity	+	a
k_{22}	R/vor	Activity	+	d
k_{23}	per/U	Productivity	+	b
k_{24}	$(U - ma)/per$	Productivity	-	c
k_{25}	U/V	Turn over	-	a, d, e
k_{26}	$EBIT/V$	Profitability	-	a, d
k_{27}	$EBIT/U$	Profitability	-	d, f
k_{28}	V/p^*	Size	-	a, e
k_{29}	U/p^*	Size	-	a, e
k_{30}	U^t/U^{t-1}	Growth rate	-/+	a
k_{31}	$(FK^t/V^t)/(FK^{t-1}/V^{t-1})$	Leverage Change	+	a
k_{32}	$(FK_B^t/v^t)/(FK_B^{t-1}/v^{t-1})$	Leverage Change	+	a

TABELLE 5.1. Klassifizierung der Kennzahlen und Hypothesenbildung. Die verwendeten Kennzahlen sind nach den Studien von (a) Falkenstein, Boral und Carty (2000), (b) Khandani, Lozano und Carty (2001), (c) Lettmayr (2001), (d) Chen und Shimerda (1981), (e) Kahaya und Theodossiou (1999), (f) Crouhy, Galai und Mark (2001) ausgewählt. Q: Hayden (2002)

Extrembereichen der latenten Variablen führen nicht mehr zu wesentlichen Änderungen der Eintrittswahrscheinlichkeit. Gleiche Veränderungen

in den Beobachtungswerten der unabhängigen Variablen X wirken sich unterschiedlich auf die Eintrittswahrscheinlichkeiten $P(y_i = 1)$ aus. Aufgrund des logistischen Funktionsverlaufs ist jedoch die Richtung des Einflusses auf die unabhängigen Variablen erkennbar.

Nachdem die Koeffizienten geschätzt wurden, stellt sich die Frage, wie signifikant die Variablen in demjenigen Modell sind. Ein möglicher Ansatz bei der Logitregression ist der Vergleich der beobachteten Werte mit denen der geschätzten, die man von Modellen erhalten hat, die sowohl mit als auch ohne der jeweiligen Variable angewendet wurden.

Anschließend werden die bereits besprochenen Konzepte der *Cumulative Accuracy Profiles* und *Accuracy Ratios* von Sobehart et al. (2000) angewendet. Dieses Modell dient dazu, sowohl die Stichhaltigkeit als auch die Richtigkeit des Kreditrisikomodells zu überprüfen.

Nachdem die Accuracy Ratios berechnet wurden, wählt man jeweils die Kennzahlen mit P-Wert $p < 0,05$ (*Marginal Significance Level*) pro Kreditrisikofaktor aus. Unter all diesen Kennzahlen (mit $p < 0,05$) wähle man jene mit maximalem *Accuracy Ratio* aus, um damit ein multivariates Modell zu gestalten.

Der P-Wert ist das Ergebnis eines Signifikanztests zur Prüfung der vorab aufgestellten Nullhypothese. Ist der P-Wert kleiner als das, ebenfalls vorab, gewählte Irrtums-(Signifikanz-)Niveau, dann gilt das Ergebnis als statistisch signifikant.

Kennzahl	AR	p	β	χ^2
k_1	49.139	0.0006761	1.145856	11.55381
k_2	34.374	0.0003408	-0.9371329	12.83152
k_3	25.978	0.120709	0.618187	6.300425
k_4	30.008	0.0129877	0.7037361	6.170785
k_5	20.465	0.2968889	0.3655879	1.088116

TABELLE 5.2. Ergebnisse der univariaten Logit-Regression: Leverage

2.1. Leverage. Der AR-Wert der Kennzahl 1 (Verschuldungsgrad) ist auffallend hoch und der p-Wert liegt unter dem (*Marginal Significance Level*). Daher wird Kennzahl 1 für das multivariate Modell ausgewählt. Kennzahl

2 (Eigenkapitalquote) wird nicht ausgewählt, da nur eine Kennzahl pro Kategorie in das Modell einfließt.

Kennzahl	AR	p	β	χ^2
k_6	28.019	0.0440396	-100.363	4.055044

TABELLE 5.3. Ergebnisse der univariaten Logit-Regression:
Debt Coverage

2.2. Debt Coverage. Die Kennzahl 6 (*EBIT/Interest Expenses*) überzeugt mit einem hohen AR-Wert und mit einem p-Wert, der unter dem (*Marginal Significance Level*) liegt. Die Kennzahl wird daher im multivariaten Modell verwendet.

Kennzahl	AR	p	β	χ^2
k_7	-1.194	0.7283645	0.0021558	0.1206186
k_8	38.157	0.380603	-0.3492755	0.7687533
k_9	1.248	0.8586928	-0.537712	0.0316977
k_{10}	4.808	0.9936649	-0.0023135	0.000063
k_{11}	14.529	0.6122624	-0.2828129	0.2568935
k_{12}	9.061	0.6846281	-0.0481893	0.1649626
k_{13}	-5.764	0.5377861	0.149974	0.3796585
k_{14}	-7.594	0.8622444	0.047421	0.0301084
k_{15}	10.66	0.5645608	-0.0801551	0.3318687
k_{16}	6534	0.6209217	-0.0602283	0.2445742

TABELLE 5.4. Ergebnisse der univariaten Logit-Regression: Liquidity

2.3. Liquidity. Die Kennzahlen aus der Gruppe Liquidity werden für das weitere Modell nicht verwendet. Der AR-Wert der Kennzahl 8 ist recht hoch, nur liegt der p-Wert über dem (*Marginal Significance Level*).

Kennzahl	AR	p	β	χ^2
k_{17}	-1.807	0.6825142	-1.976154	0.1673094
k_{18}	2.483	0.7717623	-0.2280524	0.0841409
k_{19}	42.705	0.0581539	-2.643394	3.589267
k_{20}	24.385	0.0137411	-4.556792	6.071109
k_{21}	14.486	0.0210218	-0.4625115	5.324981
k_{22}	13.833	0.0870166	-0.378558	2.928717

TABELLE 5.5. Ergebnisse der univariaten Logit-Regression: Activity Ratios

2.4. Activity Ratios. Kennzahl 20 (*Trade receivables/Liabilities*) wird aufgrund des AR-Werts und des passenden p-Werts ausgewählt. Kennzahl 19 (*Trade Receivables/Net Sales*) kommt aufgrund des p-Werts nicht in Frage.

Kennzahl	AR	p	β	χ^2
k_{23}	0.131	0.7347333	-0.2440528	0.1148098
k_{24}	6.972	0.5286781	-0.0565858	0.39669318

TABELLE 5.6. Ergebnisse der univariaten Logit-Regression: Productivity

2.5. Productivity. In dieser Kennzahlengruppe sind die AR-Werte zu gering und die Kennzahlen sind auch nicht signifikant.

Kennzahl	AR	p	β	χ^2
k_{25}	18.538	0.0560771	-0.3397574	3.649765

TABELLE 5.7. Ergebnisse der univariaten Logit-Regression: Turnover

2.6. Turnover. Aufgrund des AR-Wertes wird die Kennzahl 25 (*Netsales/Assets*) in das multivariate Modell einbezogen.

Kennzahl	AR	p	β	χ^2
k_{26}	5.678	0.6301303	-0.42074	0.2318841
k_{27}	0.151	0.7085233	-0.1966605	0.1397564

TABELLE 5.8. Ergebnisse der univariaten Logit-Regression: Profitability

2.7. Profitability. Die AR-Werte sind zu gering, die p-Werte über dem (*Marginal Significance Level*). Diese Kennzahlen werden im multivariaten Modell nicht verwendet.

Kennzahl	AR	p	β	χ^2
k_{28}	43.709	0.8503703	1.68e-08	0.0355873
k_{29}	2.139	0.9891655	-2.88e-09	0.0001844

TABELLE 5.9. Ergebnisse der univariaten Logit-Regression: Size

2.8. Size. Der AR-Wert der Kennzahl 28 (*Assets/VPI*) ist hoch, jedoch scheint die Kennzahl nicht signifikant zu sein. Daher wird sie nicht weiter verwendet.

Kennzahl	AR	p	β	χ^2
k_{30}	7.976	0.3644653	-0.3495804	0.8224483

TABELLE 5.10. Ergebnisse der univariaten Logit-Regression: Growth Rates

2.9. Growth Rates. Die Kennzahl wird im multivariaten Modell nicht verwendet, der AR-Wert ist zu gering.

Kennzahl	AR	p	β	χ^2
k_{31}	33.826	0.1793227	-0.8277336	1.803228
k_{32}	2.652	0.0720331	-0.3599022	3.236073

TABELLE 5.11. Ergebnisse der univariaten Logit-Regression: Change of Leverage

2.10. Change of Leverage. Der AR-Wert der Kennzahl 31 (*Liabilities/Assets*)/(*Liabilities/Assets*)*Last Year* ist recht hoch, nur liegt der p-Wert über dem (*Marginal Significance Level*). Daher wird diese Kennzahl nicht im multivariaten Modell verwendet.

Mit folgenden Kennzahlen wird nun ein Multivariates Modell hergeleitet:

Kennzahl	AR	<i>p</i>	β	χ^2
k_1	49.139	0.0006761	1.145.856	1.155.381
k_6	28.019	0.0440396	-100.363	4.055.044
k_{20}	24.385	0.0137411	-4.556.792	6.071.109
k_{25}	18.538	0.0560771	-0.3397574	3.649.765

TABELLE 5.12. Ergebnisse der univariaten Logit-Regression: Basismodell

3. Herleitung des Multivariaten Logitmodells

Die Kennzahlen k_1 , k_6 , k_{20} und k_{25} entsprechen den Kriterien des Basismodells. Im Anschluß daran wurden mittels Backward Selection die Variablen für das Modell ausgewählt:

Kennzahl	Coef.	Std.Err.	<i>z</i>	$P > z $	[95% Conf. interval]
k_1	1.229193	0.3730396	3.30	0.001	0.4980484 1.960337
k_6	-0.3893375	0.6559	-0.59	0.553	-1.674878 0.8962028
k_{20}	-2.221687	2.045815	-1.09	0.277	-6.231411 1.788037
k_{25}	-0.3317166	0.2119014	-1.57	0.117	-0.7470358 0.0836026
cons	-2.818708	0.4233469	-6.66	0.000	-3.648453 -1.988964

TABELLE 5.13. Multivariate Logit-Regression Schritt 1

Kennzahl	Coef.	Std.Err.	<i>z</i>	$P > z $	[95%Conf. Interval]
k_1	1.269348	0.3657979	3.47	0.001	0.5523976 1.986299
k_{20}	-2.214127	2.032987	-1.09	0.276	-6.198709 1.770456
k_{25}	-0.3626008	0.2072408	-1.75	0.080	-0.7687852 0.0435836
cons	-2.876545	0.4108883	-7.00	0.000	-3.681872 -2.071219

TABELLE 5.14. Multivariate Logit-Regression Schritt 2

Kennzahl	Coef.	Std.Err.	z	$P > z $	[95%Conf.	Interval]
k_1	1.348022	0.3571766	3.77	0.000	0.6479685	2.048075
k_{25}	-0.4567966	0.1957435	-2.33	0.020	-0.8404468	-0.0731463
cons	-2.96562	0.401044	-7.39	0.000	-3.751653	-2.179586

TABELLE 5.15. Multivariate Logit-Regression Schritt 3

Das Rating Modell basiert auf den Kennzahlen k_1 (Liabilities/Assets), und k_{25} (Netsales/Assets). Eine Erhöhung der Kennzahl k_1 bedeutet ein höheres Ausfallrisiko. Eine Verringerung der Kennzahl k_{25} bedeutet eine Verringerung des Ausfallrisikos.

Das Accuracy Ratio dieses Modells beträgt 37,351. Der relativ geringe AR Wert des Multivariaten Modells könnte sich aus der Stichhaltigkeit der Bilanzdaten in der Hotelindustrie erklären. Allerdings ist der AR Wert des Multivariaten Modells geringer als der AR Wert der Kennzahl k_1 (Liabilities/Assets). Der AR Wert der univariaten Modelle belief sich bei Kennzahl k_1 (Liabilities/Assets) auf 49,139 sowie bei Kennzahl k_{25} (Netsales/Assets) auf 18,538. Die Anzahl der Beobachtungen ist bei beiden Modellen gleich. Im Hinblick auf den geringen AR Wert der Kennzahl k_{25} und den p-Wert von 0.0560771 kann das Univariate Modell mit Kennzahl k_1 (Liabilities/Assets) empfohlen werden.

KAPITEL 6

Zusammenfassung

Zielsetzung dieser Arbeit war es, ein einfaches, praktisches und dennoch effizientes Kreditrisikomodell für den österreichischen Hotel- und Tourismussektor zu erstellen.

Zuerst wurde Kreditrisiko definiert und beschrieben, sowie sämtliche Arten von Kreditrisikomodellen aufgezählt und erläutert. Es stellte sich heraus, dass das Logit-Modell eine vorteilhafte Methode für eine Prognose der Ausfallrate ist. Weiters wurde auf die Bedeutung von Kennzahlen insbesondere für Kreditrisikomodelle eingegangen, Klassifizierungsmöglichkeiten wurden erläutert. Weiters wurde in diesem Kapitel auch auf die Definition von Ausfall eingegangen. Das *Accuracy Ratio*, als verwendete Technik für die Auswahl der Kennzahlen für das Logit-Modell wurde erklärt.

Kapitel 2 beschreibt die wichtige Positionierung der österreichischen Hotelindustrie für die österreichische Wirtschaft und zeigt, dass sie auch im weltweiten Vergleich einen namhaften Platz einnimmt. Die Struktur der österreichischen Hotelindustrie auf Basis von Klein- und Mittelbetrieben wird beschrieben.

Im Kapitel 3 wurde die Methodik und die Statistik beschrieben. Als Ergebnis der Multivariaten Logit-Regression ergab sich ein Modell aus den Kennzahlen k_1 , dh. *Liabilities/Assets* aus der Kennzahlengruppe *Liquidity* und Kennzahl k_{25} (*Netsales(Assets)*) aus der Kennzahlengruppe *Turnover*. Dadurch, dass das *Accuracy Ratio* des Multivariaten Modells geringer ist als das *Accuracy Ratio* des Univariaten Modells mit der Kennzahl k_1 (*Liabilities/Assets*), wird empfohlen, auf Kennzahl k_{25} zu verzichten und auf das Univariate Modell zurückzugreifen.

Literaturverzeichnis

- Klaus Backhaus, Bernd Erichson, Wulff Plinke und Rolf Weiber. *Multivariate Analysemethoden*. Springer, Berlin Heidelberg, 2003.
- Klaus Backhaus, Bernd Erichson, Wulff Plinke und Rolf Weiber. *Multivariate Analysemethoden*. Springer, Berlin Heidelberg, 2008.
- Basel Committee on Banking Supervision. Working Paper on the Internal Ratings-Based Approach to Specialised Lending Exposures. BCBS Working Papers No 9, 2001.
- William H. Beaver. Alternative Accounting Measures as Predictors of Failure. *Accounting Review*, 43(1):113–122, 1968.
- Hans E. Buschgen. *Bankbetriebslehre: Bankgeschäfte und Bankmanagement*. Gabler, Wiesbaden, 1999.
- John B. Caouette, Edward I. Altman und Paul Narayanan. *Managing credit risk: the next great financial challenge*. John Wiley & Sons, Inc., New York et al., 1998.
- Kung H. Chen und Thomas Shimerda. An Empirical Analysis of Useful Financial Ratios. *Financial Management*, 10(1):51–60, 1981.
- Michel Crouhy, Dan Galai und Robert Mark. A comparative analysis of current credit risk models. *Journal of Banking & Finance*, 24:59–117, 2000.
- Russel Davidson und James G. MacKinnon. *Estimation and inference in econometrics*. Oxford University Press, New York Oxford, 1993.
- Werner Doralt, Hrsg. *Internationale Rechnungslegung IAS/IFRS 2009*. Linde, Wien, 2009.

- Robert O. Edmister. An Empirical Test of Financial Ratio Analysis for Small Business Failure. *Journal of Financial and Quantitative Analysis*, 7:1477 – 1493, 1972.
- Bernd Engelmann, Evelyn Hayden und Dirk Tasche. Testing Rating Accuracy. *Risk*, 20, 2003.
- Gunter Franke und Herbert Hax. *Finanzwirtschaft des Unternehmens und Kapitalmarkt*. Springer, Berlin Heidelberg, 2004.
- Evelyn Hayden. *Modeling an Accounting-Based Rating System for Austrian Firms*. Dissertation, Universität Wien, Fakultät für Wirtschaftswissenschaften und Informatik, 2002.
- David W. Hosmer und Stanley Lemeshow. *Applied Logistic Regression*. John Wiley & Sons, Inc., New York et al., 1989.
- Hans-Georg Kantner. Insolvenzstatistik 2008. KSV1860 Onlinepublikation, 2009a. URL <http://www.ksv.at/KSV/1870/de/pdf/statistik/InsolvenzstatistikUnternehmen2008.pdf>.
- Hans-Georg Kantner. Insolvenzursachen 2008. KSV1860 Onlinepublikation, 2009b. URL http://www.ksv.at/KSV/1870/de/5presse/2medi enarchiv/0pressemeldungen/2009-04/insolvenzursachen2008_hgk/1240903160732_090423_Insolvenzursachen_Kantner.pdf.
- Matthias Kerling und Thorsten Poddig. Klassifikation von Unternehmen mittels KNN. In Heinz Rehkugler, Hrsg., *Neuronale Netze in der Ökonomie: Grundlagen und finanzwirtschaftliche Anwendungen*. Vahlen, München, 1994.
- David G. Kleinbaum und Mitchel Klein. *Logistic Regression*. Springer, New York et al., 2010.
- Kreditschutzverband. KSV-Risikobewertung Details und Hintergründe. KSV1860 Onlinepublikation, 2008. URL <http://www.ksv.at/KSV/1870/de/pdf/105RatingDetails.pdf>.

- Peter Laimer und Christa Schischeg. *Tourismus in Österreich 2008*. Verlag Österreich GmbH, Wien, 2009.
- Erkki K. Laitinen, Rolf Leppanen und Edward I. Altman. *High tech company rating: financial and operative evaluation*. National High Technology Mentor Programme, 1997.
- Robert Libby. Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence. *Journal of Accounting Research*, 13(1):150–161, 1975.
- Daniel Martin. Early warning of bank failure: A logit regression approach. *Journal of Banking and Finance*, 1(3):249–276, 1977.
- Birgit Ott. *Interne Kreditrisikomodelle*. Uhlenbuch Verlag, Bad Soden/Ts, 2001.
- Harlan D. Platt und Marjorie B. Platt. A note on the industry-relative ratios in bankruptcy prediction. *Journal of Banking and Finance*, 15(6):1183–1194, 1991.
- Peter Reichling, Daniela Bietke und Antje Henne. *Praxishandbuch Risikomanagement und Rating*. Gabler, Wiesbaden, 2007.
- Roman Rohatschek und Helmut Maukner. *Rechnungslegung nach International Financial Reporting Standards*. Manz, Wien, 2008.
- Gerd Ronning. *Mikroökonomie*. Springer, Berlin Heidelberg, 1991.
- Timo Salmi und Teppo Martikainen. A Review of the Theoretical and Empirical Basis of Financial Ratio Analysis. *The Finnish Journal of Business Economics*, 4:426–448, 1994.
- Anthony Saunders. *Financial Institutions Management: A Modern Perspective*. Irwin, Chicago et al., 1997.
- Andreas Schmidt-von Rhein und Heinz Rehkugler. KNN zur Kreditwürdigkeitsprüfung bei Privatkundenkrediten. In Heinz Rehkugler, Hrsg., *Neuronale Netze in der Ökonomie: Grundlagen und finanzwirtschaftliche Anwendungen*. Vahlen, München, 1994.

- Jorge Sobehart und Sean Keenan. Measuring default accurately. *Credit Risk Special Report*, pages 31 – 33, 2001.
- Jorge Sobehart, Sean Keenan und Roger Stein. Validation methodologies for default risk models. *Credit*, pages 51–56, May 2000.
- Statistik Austria. Tourismus im Überblick. Onlinedienst der Statistik Austria, 2012. URL http://www.statistik.at/web_de/services/stat_uebersichten/tourismus/index.html.
- Statistik Austria. Leistung- und Strukturstatistik 2007. Onlinedienst der Statistik Austria, 2009a. URL http://www.statistik.at/web_de/statistiken/produktion_und_bauwesen/leistungs_und_strukturdaten/index.html.
- Statistik Austria. Internationaler Reiseverkehr: Einnahmen und Ausgaben 1996 bis 2008. Onlinedienst der Statistik Austria, 2009b. URL http://www.statistik.at/web_de/statistiken/tourismus/reiseverkehrsbilanz/index.html.
- Alex Stomper und Christine Zulehner. Why Leverage Distorts Investment. Working Paper University of Vienna, March 2004.
- Joachim Hartung und Bärbel Elpelt. *Multivariate Statistik: Lehr- und Handbuch der angewandten Statistik*. Oldenbourg, München, 1989.
- Robert Craig West. A Factor Analytic Approach to Bank Condition. *Journal of Banking and Finance*, 9(2):253–266, 1985.
- Hans Georg Zimmermann. Neuronale Netze als Entscheidungskalkül. In Heinz Rehkugler, Hrsg., *Neuronale Netze in der Ökonomie: Grundlagen und finanzwirtschaftliche Anwendungen*. Vahlen, München, 1994.

Abstract

Aim of this paper is to develop a credit risk model for the Austrian Hospitality Sector. The research question includes whether it is useful to develop a industry specific credit risk model and furthermore which techniques are useful to develop a credit risk model based on the specifications of the industry. The developed model shall enable to predict the default probability as accurate as possible using a multivariate statistic model. The credit risk model is based on a logit regression using a data pool provided primarily by the Austrian Hotel and Tourism Bank as well as the Austrian National Bank, the Austrian Institute of Small Business Research and 3 major commercial Austrian banks. Firstly the parameters are estimated by univariate regression. The multivariate model is developed focussing on the significance using back ward selection. The performance of the estimated models are evaluated by the concepts of cumulative accuracy profiles and accuracy ratios. Results of the multivariate model are the significant accounting ratios Liabilities/Assets and Netsales/Assets. This can be explained by very high leverage being a specification of this industry. Interesting and surprising is that other accounting ratios do not seem to be significant. The accuracy ratio-value of the univariate model is higher than the accuracy ratio-value of the multivariate model, which is paradox. Further research prospects can be to examine this contradiction, include additional criteria for the selection of variables for the multivariate model and to extend the model to a multinomial logit-regression model.

Abstract

Zielsetzung der Arbeit ist die Erstellung eines Kreditrisikomodells für die österreichische Hotel- und Tourismusbranche. Die Forschungsfrage umfasst die Frage ob es sinnvoll ist ein branchenspezifisches Kreditrisikomodell zu entwickeln, darauf folgend, welche Methoden zur Erstellung dieses Modells unter Berücksichtigung der spezifische Eigenheiten der Branche in Frage kommen. Es soll ein Modell entwickelt werden, anhand dessen eine sichere Prognose der Ausfallswahrscheinlichkeit möglich ist. Der Ansatz ist ein multivariates statistisches Modell. Das Kreditrisikomodell wird mittels Logit-Regression auf Basis eines Datensatzes der Österreichischen Hotel- und Tourismusbank bzw. der ÖNB, KMU Forschung Austria und drei weiteren österreichischen Banken erstellt, wobei zuerst die Parameter mittels univariater Regression geschätzt werden und danach das multivariate Modell auf Basis der Signifikanz mittels Backward Selection gebildet wird. Die Richtigkeit des Modells wird mit dem Konzept der Cumulative Accuracy Profiles und des Accuracy Ratios überprüft. Im Ergebnis des multivariaten Modells zeigt sich, dass die Kennzahlen Liabilities/Assets, also der Verschuldungsgrad und Net-sales/Assets, die Umschlagshäufigkeit des Gesamtkapitals, besonders signifikant sind. Das ist darauf zurückzuführen, dass der Verschuldungsgrad der Branche sehr hoch ist. Trotzdem überrascht, dass sonstige branchenspezifische Kennzahlen keine Rolle zu spielen scheinen. Paradox ist, dass der Accuracy Ratio-Wert der Kennzahl Liabilities/Assets, des univariaten Modells höher ist, als der Accuracy Ratio-Wert des multivariaten Modells. Weitere Forschungsimpulse können diesen augenscheinlichen Widerspruch untersuchen, weitere Kriterien zur Auswahl der Variablen für das multivariate Modell einbeziehen bzw. das Modell auf ein multinomiales Logit-Regressionsmodell erweitern.

Über die Autorin

Judith Spiegl, geb. 1973 in Wien, Matura 1992 an der *Höheren Bundeslehranstalt für wirtschaftliche Berufe* in Baden, von 1992-1996 McKinsey & Comp. Inc., von 2004-2007 Forschungsförderungsgesellschaft.

Ab 1994 Studium der Internationalen Betriebswirtschaft an der *Universität Wien* mit Auslandsjahr am *BI, Handelshøyskolen*, Norwegen, Schwerpunkt Strategy & Finance.

Judith Spiegl hat zwei Kinder und lebt in Baden.

