

Kaiser, Ulrich; Szczesny, Andrea

Working Paper

Einfache ökonometrische Verfahren für die Kreditrisikomessung

CoFE Discussion Paper, No. 00/28

Provided in Cooperation with:

University of Konstanz, Center of Finance and Econometrics (CoFE)

Suggested Citation: Kaiser, Ulrich; Szczesny, Andrea (2000) : Einfache ökonometrische Verfahren für die Kreditrisikomessung, CoFE Discussion Paper, No. 00/28, University of Konstanz, Center of Finance and Econometrics (CoFE), Konstanz, <https://nbn-resolving.de/urn:nbn:de:bsz:352-opus-5775>

This Version is available at:

<https://hdl.handle.net/10419/85213>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Einfache ökonometrische Verfahren für die Kreditrisikomessung*

Ulrich Kaiser und Andrea Szczesny

Zusammenfassung

Dieser Beitrag stellt verschiedene ökonometrische Methoden zur Bewertung und Berechnung von Kreditausfallrisiken vor und wendet diese auf einen aus Kreditakten von sechs deutschen Universalbanken zusammengestellten Datensatz an. Im Mittelpunkt stehen dabei (i) binäre bzw. geordnete Logit- und Probitmodelle, mit deren Hilfe die Ausfallwahrscheinlichkeit eines Kredites geschätzt werden kann, sowie (ii) Verweildauermodelle, mit denen die Gefahr eines Kreditausfalls unter Berücksichtigung der Verweildauer im Nichtausfallzustand quantifiziert werden kann. Beispiele und Interpretationshilfen zu den jeweils vorgestellten Methoden erleichtern den Zugang zu diesen Modellen. Es werden zahlreiche Hinweise auf weiterführende Literatur gegeben.

Abstract

This paper describes simple econometric methods for the analysis of credit risk and applies them to a data set obtained from credit files taken from six large German universal banks. The paper focuses on (i) binary and ordered probit/logit models which enable the credit analyst to quantify the default probability of an individual credit, and (ii) on duration models capable of estimating the default probability of a credit at a certain point in time given that there was no default until then. Empirical examples for the methods facilitate the understanding of the econometric models described in the paper. Numerous suggestions for further reading complete this short walk down the econometric quantification of credit risk.

* Dieser Beitrag hat von hilfreichen Kommentaren von Bernhard Bookmann, Christian Ernst und Bettina Peters sowie vor allem von François Laisney stark profitiert. Ulrich Kaiser dankt der Deutschen Forschungsgemeinschaft für finanzielle Unterstützung im Rahmen des Schwerpunktprogramms „Industrieökonomik und Inputmärkte“ (Projekt PF331/3-3). Beide Autoren sind den am Projekt „Kreditmanagement“ des Center for Financial Studies (CFS) beteiligten Banken und dem CFS für die Bereitstellung der Daten zu Dank verpflichtet.

1 Einführung

Das Kreditgeschäft als traditionelles Betätigungsfeld der Banken steht unter Reformdruck. Vor dem Hintergrund der gestiegenen Anforderungen an das Risikomanagement der Banken hat auch der „Basler Ausschuss für Bankenaufsicht“, ein von den G-10 Staaten im Jahre 1974 ins Leben gerufenes Forum zur internationalen Koordination der Bankenaufsicht, Reformvorschläge zum Risikomanagement formuliert. Im Juli 1999 wurde vom Basler Ausschuss ein Konsultationspapier mit Vorschlägen zur Neuregelung der angemessenen Eigenkapitalausstattung für Banken (Basel II) vorgelegt (Basel Committee on Banking Supervision, 1999). Diese sollen den Basler Akkord von 1988 (Basel I) ersetzen (Basel Committee on Banking Supervision, 1988). Basel II bringt vor allem Reformen für das Unternehmenskreditgeschäft mit sich, das bislang einer Pauschalhinterlegung von acht Prozent des Eigenkapitals unterlag. Revolutionär ist vor allem der Vorschlag, dass Banken ihre internen Bonitätseinstufungen nutzen dürfen, um eine risikoadäquate Eigenkapitalunterlegung zu erreichen (Internal Ratings-Based Approach, IRB). Der Basler Ausschuss stellt in einer im Frühjahr 1999 durchgeführten Studie unter rund 30 Banken der G-10-Staaten zur gängigen Praxis fest, dass sich die Methoden der Risikomessung im Kreditgeschäft erheblich unterscheiden. Die Bandbreite reicht von vollständig auf Expertenaussagen basierenden Systemen bis hin zu vollständig auf statistischen Methoden aufbauenden Methoden. Zudem stellte der Ausschuss fest, dass es noch erheblich an Datenmaterial mangle. Sowohl bei der bankinternen Quantifizierung von Risiken und erst recht bei der erforderlichen Validierung der Rating-Systeme durch die Bankenaufsicht werden allerdings ausreichend Daten gebraucht (Basel Committee on Banking Supervision, 2000a).

Der Ausschuss legt in seinen geplanten Reformen besonderen Wert darauf, den Banken Anreize zu setzen, ihre Methoden zu Risikomessung und -management weiterzuentwickeln. Der Kreditrisikobereich ist, was die angewendete Methodik betrifft, im Vergleich zum Marktrisikobereich weit zurück geblieben. Üblicherweise werden Kreditmerkmale, die sich in der Vergangenheit als risikobestimmend herausgestellt haben, entsprechend gesammelter Erfahrungswerte gewichtet und zu einem Kredit-Scoring verdichtet. Zur Bestimmung der Faktoren und Gewichte werden teilweise einfache statistische Verfahren eingesetzt. Ein Vorteil dieser Vorgehensweise ist sicherlich der geringe Aufwand bei der Konstruktion des Systems und die anschließend einfache Anwendbarkeit. Diese Vorgehensweise mag zwar in der Vergangenheit durchaus angemessen gewesen sein. Es zeichnet sich jedoch ab, dass das Kreditrisikomanagement zu einem der entscheidenden Wettbewerbsfaktoren der Finanzindustrie werden wird. Ein modernes Kreditrisikomanagement wird in Zukunft mit höheren methodischen Anforderungen verbunden sein.

Über die bislang angewendeten Verfahren hinaus, die mehr oder minder auf Expertenwissen beruhen, gibt es eine Reihe einfacher ökonomischer Metho-

den, die bei der Messung von Ausfallrisiken gute Dienste leisten können. Mit ihrer Hilfe können aus dem Datenmaterial der Banken – also aus den in der Vergangenheit gesammelten Erfahrungen – wertvolle Erkenntnisse gewonnen und für das Risikomanagement nutzbar gemacht werden. Die Verfahren sind in Standardsoftwarepaketen verfügbar und einfach anwendbar. Trotzdem kommen sie bislang in der Praxis nicht oder nur selten zum Einsatz. Ein Grund dafür liegt wahrscheinlich darin, dass sie bislang in anderen Fachgebieten genutzt wurden, beispielsweise in der Arbeitsmarkt- und Industrieökonomik sowie in den Sozialwissenschaften. Das Ziel dieser Arbeit liegt darin, den Zugang zu den Methoden zu erleichtern, indem ihre wesentlichen Eigenschaften im Kontext der Risikomessung anschaulich beschrieben. Darüber hinaus werden zahlreiche Interpretationshilfen anhand von Schätzungen aus dem Bereich der Kreditrisikomessung gegeben.

Die Kreditrisikomessungsmethoden, die in diesem Aufsatz beschrieben werden, haben eines gemeinsam: Sie modellieren Variablen, die qualitative Ausprägungen besitzen. So kann ein Kredit verschiedene Qualitätszustände annehmen. Er kann z. B. vom Kreditnehmer vollständig bedient werden, nur teilweise bedient werden oder ausfallen. Interessiert den Analysten lediglich die Wahrscheinlichkeit eines Kreditausfalles gegenüber der Wahrscheinlichkeit der fortlaufenden Bedienung eines Kredites, so stellen *binäre Probit- oder Logitmodelle* einen geeigneten Analyseansatz dar. Diese Modelle werden in Abschnitt 3.1 vorgestellt. Solche Zweizustandsmodelle sind natürlich dann unzureichend, wenn der Analyst verschiedene Zustände beobachtet, die ein Kredit annimmt. Folgen diese Zustände einer „natürlichen Ordnung“, so können sie mit *geordneten Probit/Logitmodellen* (Abschnitt 3.2) untersucht werden. Beobachtet der Analyst z. B. neben den beiden Zuständen „Ausfall“ und „Bedienung“ noch den Zustand „teilweise Bedienung“, so folgen die Zustände einer natürlichen Ordnung und die Determinanten dieser Zustände können mit solchen geordneten Probit-/Logitmodellen abgebildet werden.

Ein Analyst wird zunächst daran interessiert sein, die Robustheit seiner Kreditausfallsschätzungen zu überprüfen und einen Indikator für die Güte seines Modells zu finden. In Abschnitt 3.3 stellen wir Spezifikationstests und Gütemaße vor.

Auf Erweiterungsmöglichkeiten der binären und geordneten Probit- bzw. Logitmodelle wird in Abschnitt 3.4 eingegangen. Dabei werden Paneldatenmodelle sowie Mehrgleichungsmodelle mit Simultanität und mit Korrelation der Fehlerterme skizziert.

Können die verschiedenen Zustände, in denen sich ein Kredit befindet bzw. befunden hat, über einen längeren Zeitraum beobachtet werden, so ist es interessant, die Wahrscheinlichkeit dafür zu berechnen, dass ein Kredit von einem Zustand in den anderen übergeht. In Abschnitt 4 stellen wir deshalb das Verfahren vor, mit dem solche Übergangsraten berechnet werden können. Mit Markovketten (Unterabschnitt 4.1) können Wahrscheinlichkeiten für die Übergänge von

einem Zustand in den anderen modelliert werden. Ist man daran interessiert, die Wahrscheinlichkeiten für den Übergang in einen anderen Zustand zu berechnen, *gegeben*, dass ein Kredit eine bestimmte Periode in einem Zustand verweilt hat, dann ist der Kaplan-Meier Schätzer ein geeignetes Analyseinstrument. Er wird in Abschnitt 4.2 beschrieben. Sowohl Markovketten als auch Verweildauermodelle modellieren direkt die Zeitdimension eines Datensatzes, der Informationen über Kreditnehmer zu verschiedenen Zeitpunkten enthält. Während Markovketten jedoch die Wahrscheinlichkeit eines Übergangs von einem Zustand in den anderen abbilden, ohne dass die Verweildauer in einem Zustand mitberücksichtigt wird, kann die Zeitdimension mit Hilfe des Kaplan-Meier-Schätzers explizit modelliert werden. Hier stellt sich die Frage: Gegeben, dass ein Kredit nach 2 Jahren noch nicht ausgefallen ist, wie hoch ist die Wahrscheinlichkeit, dass er auch weiterhin nicht ausfällt?

Sowohl der Kaplan-Meier-Schätzer als auch einfache Markovketten berücksichtigen jedoch keine Variablen, die den Übergang von einem Zustand in den anderen erklären können.¹ Der Heterogenität von Kreditnehmern wird mit diesen nichtparametrischen Modellen keine Rechnung getragen. Modelle, die erklärende Variablen in der Schätzung des Kreditausfallrisikos – des „hazard“ – berücksichtigen, sind die Hazardratenmodelle, die in Abschnitt 5 dargestellt werden. Mit Hilfe von Hazardratenmodellen kann das Risiko eines Kreditausfalls unter der Berücksichtigung von Charakteristika des Kreditnehmers berechnet werden. Dem Kreditanalysten stehen dabei zwei Schätzverfahren zur Verfügung, das Modell der proportionalen Hazardrate (Abschnitt 5.1) und der Partial-Likelihood-Ansatz (Abschnitt 5.2). In Abschnitt 5.3 wird auf weiterführende Literatur verwiesen, die z. B. die Zeitabhängigkeit erklärender Variablen einbezieht und auf andere Erweiterungsmöglichkeiten der hier vorgestellten Grundmodelle eingeht.

In Abschnitt 6 werden mikroökonomische Methoden skizziert, die im Rahmen dieses Aufsatzes nicht besprochen werden konnten, aber für spezielle Fragen des Kreditrisikomanagements dennoch von Bedeutung sein können.

Abschnitt 7 fasst die in diesem Beitrag vorgestellten Analysemethoden und Ergebnisse zusammen.

Für eine Vielzahl der im folgenden vorgestellten Analysemethoden werden empirische Beispiele auf der Grundlage eines Kreditakten-Datensatzes gegeben. Dieser Datensatz ist in Abschnitt 2 genauer beschrieben.

Sämtliche hier vorgestellten Modelle sind in gängiger Standardsoftware implementiert. Für mikroökonomische Methoden empfehlen sich aufgrund der Vielzahl von Applikationen die Programme STATA und Limdep.² Die in diesem

1 In einer aktuellen Arbeit lassen Nguyen Van et al. (2000) auch erklärende Variablen in Markovkettenmodellen zu.

2 Informationen zu STATA sind im Internet unter <http://www.stata.com> zu finden, Informationen zu Limdep können unter <http://www.limdep.com> abgerufen werden.

Beitrag durchgeführten Schätzungen wurden mit STATA 6.0 durchgeführt, alle hier vorgestellten Modelle sind aber auch in Limdep enthalten.

2 *Daten*

2.1 *Datenquelle*

Alle Schätzungen werden auf der Basis eines Datensatzes durchgeführt, der im Rahmen des Projekts „Kreditmanagement“ des Instituts für Kapitalmarktforschung (Center for Financial Studies, CFS), erhoben wurde. Es handelt sich dabei um Informationen aus 260 Kreditakten mittelständischer Unternehmen der Kreditinstitute Bayerische Vereinsbank, Commerzbank, Deutsche Bank, Deutsche Genossenschaftsbank, Dresdner Bank und Westdeutsche Landesbank für den Zeitraum von 1992 bis 1998 (nähere Angaben finden sich unter anderem in Elsas et al., 1998).³ Um Verzerrungen der Stichprobe hinsichtlich problembehafteter bzw. nicht problembehafteter Kredite zu vermeiden, wurde der Datensatz hinsichtlich dieses Kriteriums geschichtet.⁴ Für die hier durchgeführten Untersuchungen standen uns Daten von jeweils rund einhundert Kreditkunden zur Verfügung, bei denen es während des Beobachtungszeitraums zu Problemen gekommen ist bzw. bei denen keine Rückzahlungsschwierigkeiten verzeichnet wurden.

Im folgenden Abschnitt werden die Begriffe Ausfall, Teilausfall und Vollausfall definiert. Darüber hinaus wird ein Überblick über die in den einfachen Modellen verwendeten Variablen gegeben.

2.2 *Definitionen*

Ausfall:

Um die Beschreibung der methodischen Verfahren übersichtlich zu gestalten, sprechen wir durchgehend von Ausfall und Ausfallwahrscheinlichkeit. Unter dem Ereignis eines Ausfalls subsumieren wir betriebswirtschaftlich gesehen allerdings nicht nur Vollausfälle von Krediten, sondern das Auftreten jeglicher Schwierigkeiten, die bei der Vertragserfüllung auftreten können, da sie mit zusätzlichen Kosten für die Bank verbunden sind. Dazu gehören die Stundung von Zins- und Tilgungszahlungen, das Einfordern zusätzlicher Sicherheiten, das Einleiten von Umstrukturierungsmaßnahmen im operativen Geschäft der Unterneh-

3 Weitere auf dem Datensatz aufbauende empirische Studien finden sich auf der Homepage des CFS (http://www.ifk-cfs.de/pages/veroeff/cfswor/index_d.htm).

4 Auf Probleme von Stichprobenselektionsverzerrungen und deren Korrektur gehen wir in Abschnitt 3.4.2 ein.

men, das Verwerten von Sicherheiten, die Fälligstellung von Krediten, Abwicklungen, Vergleiche, Konkurse und Sanierungen.

Kein Ausfall, Teilausfall und Vollausfall:

Wenn eine Unterteilung der Daten in die drei Kategorien kein Ausfall, Teilausfall und Vollausfall vorgenommen wird, dann verstehen wir unter Teilausfällen das Auftauchen von Problemen bei der Erfüllung von Kreditverträgen, die zu Verlusten auf Seiten der Bank führen, aber nicht zwangsläufig in einen Vollausfall münden. Dazu gehören Probleme, die beispielsweise zu einer Stundung von Zins- und Tilgungszahlungen, zum Einfordern zusätzlicher Sicherheiten oder zur Verwertung von Sicherheiten führen. Als Vollausfall bezeichnen wir schwere Probleme bei der Erfüllung des Kreditvertrages, die zum Beispiel eine Fälligkeit nach sich ziehen sowie Abwicklungen, Vergleiche und Konkurse. Aber auch Sanierungen, die im Erfolgsfall natürlich keinen Vollausfall des Kredites bedeuten, werden zu der Kategorie „Vollausfall“ gezählt.

2.3 *Verwendete Variablen*

Dieser Beitrag stellt die Anwendung unterschiedlicher Verfahren aus der Ökonometrie in den Mittelpunkt. Daher wurden die Modelle des Anwendungsbeispiels aus der Kreditrisikomessung bewusst einfach strukturiert.

Wir berücksichtigen Informationen zur Unternehmensgröße (Umsatz), zur Rechtsform (Haftungsbeschränkung) und zur Branchenzugehörigkeit der Unternehmen. Kennzahlen (Eigenkapitalquote, dynamischer Cash-Flow und Anlagendeckung) geben Informationen über die Vermögens-, Finanz- und Ertragslage der Unternehmen. Eventuelle gesamtwirtschaftliche Einflüsse werden mit Hilfe von Indikatorvariablen für die unterschiedlichen Beobachtungszeitpunkte berücksichtigt. Die Literatur zu möglichen Indikatoren für das Kreditrisiko und dessen Modellierung ist kaum zu überschauen. Einen guten Einstieg bieten Altman und Saunders (1998) sowie die Veröffentlichungen des Basler Ausschusses für Bankenaufsicht (2000a und 2000b). Die folgenden Variablen werden verwendet:

- *Ausfall*: Binärvariable, die bei Problemen mit der Vertragserfüllung den Wert 1 annimmt und sonst den Wert 0 trägt (siehe dazu auch Kapitel 2.2).
- *Ausfall_3*: Variable, die in drei Abstufungen Probleme anzeigt. Ein Wert von 0 besagt, dass keine Probleme vorliegen, ein Wert von 1 zeigt Probleme an, die noch keinen Totalausfall des Kredits bedeuten und ein Wert von 2 steht für schwere Probleme wie Abwicklung, Vergleich oder Konkurs (siehe dazu auch Kapitel 2.2).
- $\ln(\text{Umsatz})$: Variable, welche die Größe des Unternehmens anhand der Höhe seiner Umsätze repräsentiert. Dabei werden die Umsätze mit Hilfe des natürlichen Logarithmus transformiert.
- $\ln(\text{Umsatz})^2$: Um eventuelle nichtlineare Einflüsse der Unternehmensgröße zu berücksichtigen, werden quadrierte logarithmierte Umsätze in die Schätzungen aufgenommen.

- Eigenkapitalquote: Eigenkapitalquote des Unternehmens, berechnet als Quotient aus Eigenkapital und Bilanzsumme.
- Cash Flow: Dynamischer Cash Flow, berechnet als Quotient aus Cash Flow und Nettoverbindlichkeiten des Unternehmens.
- Anlagendeckungsgrad: Anlagendeckungsgrad, berechnet als Quotient aus mittel- und langfristiger Passiva und mittel- und langfristiger Aktiva.
- Beschr. Haftung: Binärvariable, die den Wert 1 annimmt, wenn die Unternehmer nur beschränkt haften, ansonsten den Wert 0 trägt.
- 1992, 1993, ..., 1998: Binärvariablen, die anzeigen, aus welchem Jahr die Beobachtung stammt, wobei das Jahr 1992 in den Schätzungen als Referenz genommen wird.
- Maschinenbau: Binärvariable, die Unternehmen aus dem Sektor Maschinenbau kennzeichnet, wird in den Schätzungen als Referenzgröße genutzt.
- Verarb. Gew.: Binärvariable, die Unternehmen aus dem übrigen Sektor des verarbeitenden Gewerbes kennzeichnet.
- Baugewerbe: Binärvariable für Unternehmen aus dem Baugewerbe.
- Handel: Binärvariable für Unternehmen aus dem Bereich Handel
- Sonstige: Binärvariable für sonstige Unternehmen, die zum größten Teil aus dem Dienstleistungsbereich und dem Bereich Transport und Logistik stammen.

3 Modelle für qualitative abhängige Variablen

3.1 Zweizustandsmodelle: Binäre Logit- und Probitmodelle

Der einfachste Ansatz zur Modellierung eines Kreditausfallrisikos ist das binäre Probitmodell. In diesem einfachsten Fall können zwei mögliche Zustände eines Kredites beobachtet werden: er wird bedient oder er fällt aus. Aus dieser Information wird eine abhängige Variable konstruiert, die den Wert 0 annimmt, wenn Kredit i bedient wird und die den Wert 1 annimmt, wenn er ausfällt.⁵ Diese Variable wird im folgenden mit $Ausfall_i$ abgekürzt. Das Subskript i indiziert dabei den i ten Kredit. Wir nehmen an, dass das Ausfallrisiko von den Variablen Firmengröße ($\ln(\text{Umsatz})$, $\ln(\text{Umsatz})^2$), Vermögens-, Finanz- und Ertragslage (Eigenkapitalquote, Cash-Flow, Anlagendeckungsgrad), Branchenzugehörigkeit: Verarbeitendes Gewerbe, Bauwirtschaft, Handel und Maschinenbau (als Basiskategorie) sowie der konjunkturellen Lage, die durch Zeitdummies 1993 (Jahr 1993) bis 1998 (Jahr 1998) abgebildet wird, bestimmt wird. Die Basiskategorie bildet dabei das Jahr 1992. Diese Bestimmungsgrößen werden, zusätzlich zu einer Konstanten, in einem Vektor erklärender Variablen, x_i , zusammengefasst.

5 Die Kodierung dieser Variablen ist dabei willkürlich. Die hier angegebene Spezifikation modelliert die Wahrscheinlichkeit, dass ein Kredit ausfällt. Lautete die Kodierung umgekehrt, so würde die Wahrscheinlichkeit des Nicht-Ausfalls spezifiziert.

Beim binären Probitmodell wird davon ausgegangen, dass die zu erklärende Variable dann den Wert 1 annimmt, wenn eine unbeobachtbare Variable eine bestimmte Schwelle s überschreitet, die für alle Kredite i identisch ist. Diese latente Variable setzt sich aus dem Vektor der erklärenden Variablen und einem Vektor von zu schätzenden Parametern, sowie einem unabhängig und identisch normalverteilten Zufallsterm ε_i zusammen. Wenn dieser Zufallsterm logistisch verteilt ist, ergibt sich das Logitmodell. Im Fall der Kreditausfallrisikomessung kann die unbeobachtbare, „latente“, Variable als eine gewichtete Summe von Faktoren aufgefasst werden, die letztlich zum Kreditausfall führen. Überschreitet diese Summe die Schwelle s , so kommt es zum Kreditausfall:

$$(1) \quad \text{Ausfall}_i = \begin{cases} 1 & \text{wenn } \text{Ausfall}_i^* = x_i \beta + \varepsilon_i > s \\ 0 & \text{sonst.} \end{cases}$$

Sofern der Fehlerterm ε_i einer symmetrischen Verteilung folgt, kann Gleichung (1) in folgenden Wahrscheinlichkeiten ausgedrückt werden:

$$(2) \quad \begin{aligned} P(\text{Ausfall}_i = 1 | x_i) &= P\left(\frac{\varepsilon_i > s - x_i \beta}{\sigma} = 1 - F\left(\frac{\varepsilon_i \leq s - x_i \beta}{\sigma}\right) \\ P(\text{Ausfall}_i = 0 | x_i) &= P\left(\frac{\varepsilon_i \leq s - x_i \beta}{\sigma}\right) = F\left(\frac{\varepsilon_i \leq s - x_i \beta}{\sigma}\right), \end{aligned}$$

wobei F im Logitfall die Verteilungsfunktion der logistischen Verteilung und im Probitfall die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Der Parameter σ bezeichnet die Standardabweichung des Fehlerterms. Die Formulierung in (2) macht deutlich, dass es hier um konditionale Wahrscheinlichkeiten geht. Die Wahrscheinlichkeit eines Kreditausfalls wird auf den Vektor der Unternehmenscharakteristika, x_i , bedingt. Um die Identifikation des Modells zu gewährleisten, müssen Restriktionen gesetzt werden. Standardsoftwareprogramme wie z. B. STATA setzen $\sigma = \pi / \sqrt{3}$ und $s = 0$ für das Logitmodell sowie $\sigma = 1$ und $s = 0$ für das Probitmodell. Diese Normierungen ändern an der Interpretierbarkeit des Modells nichts.⁶ Soweit es im Text nicht anders vermerkt ist, beziehen sich nun folgenden Ausführungen auf den Probitfall.

Die Berechnung des Koeffizientenvektors β erfolgt mit der Maximum-Likelihood-Methode. Dabei geht es darum, über die Wahl von β die Wahrscheinlichkeit zu maximieren, mit der der vorliegende Datensatz reproduziert wird.⁷

Die Schätzung von Paneldatenmodellen (Abschnitt 3.4.1) verlangt die Equidistanz, also einen zeitlich gleichen Abstand zwischen den einzelnen Beobach-

6 Alternativ zu dieser Normierung hätten auch folgende Restriktionen gewählt werden können (1) $\sigma = 1$ und Koeffizient der Konstanten = 0 (Mitschätzen von s) sowie (2) Koeffizient der Konstanten = 0 und $s =$ irgend ein willkürlicher Wert (Mitschätzen von σ).

7 Verbeek (2000, Kap. 6) bietet einen hervorragenden Einstieg in die Maximum-Likelihood-Methoden und in die Prinzipien numerischer Optimierung.

tungszeitpunkten. Dies ist im uns vorliegenden Datensatz nicht der Fall. Vielmehr existieren zu einigen Krediten mehr als eine Beobachtung innerhalb eines Jahres. Die Equidistanz der Beobachtungen wird deshalb dadurch hergestellt, dass lediglich die jeweils letzten Informationen zu einzelnen Krediten betrachtet werden. Ereignisse, die zwischen zwei Beobachtungszeitpunkten stattgefunden haben, werden pro Jahr kumuliert und mit der letzten Beobachtung im Jahr angegeben. Bei den in den Abschnitten 4.2 und 5 vorgestellten Modellen wird wiederum der gesamte Datensatz betrachtet, da Equidistanz hier nicht benötigt wird. Im Gegenteil: Hier geht es gerade um die Analyse der Abstände zwischen Risikozuständen.

Zur Illustration zeigt Tabelle 1 die Ergebnisse des oben beschriebenen Modells zur Schätzung der Kreditausfallwahrscheinlichkeit. Die Tabelle zeigt in den Spalten von links nach rechts die Variablenbezeichnung, den Wert des geschätzten Koeffizienten, den zugehörigen Standardfehler und das empirische Signifikanzniveau. Zusätzlich werden noch Wald-Tests auf gemeinsame Signifikanz der Umsatzvariablen, der Zeitdummies und der Branchendummies ausgewiesen

Die Schätzung des einfachen Modells deutet darauf hin, dass die Unternehmensgröße keinen signifikanten Einfluss auf die Ausfallwahrscheinlichkeit eines Kreditkunden hat. Dieser wäre beispielsweise durch Diversifikationseffekte in der Produktpalette oder eine höhere Marktmacht zu erwarten gewesen. Möglicherweise wären hier andere Größenindikatoren wie etwa die Bilanzsumme oder die Zahl der Beschäftigten besser geeignet gewesen. Die Kennzahlen zur Vermögens-, Finanz- und Ertragslage des Unternehmens zeigen eine signifikante Wirkung auf die Ausfallwahrscheinlichkeit. Die Eigenkapitalquote ist negativ und signifikant auf dem 1-Prozent-Niveau: Je höher die Eigenkapitalquote ist, desto niedriger ist die Ausfallwahrscheinlichkeit. Der Cash Flow und die Anlagendeckung sind auf dem 5-Prozent-Niveau signifikant, auch sie haben einen negativen Einfluss. Eine Haftungsbeschränkung hingegen scheint nicht mit einer signifikant höheren oder niedrigeren Ausfallwahrscheinlichkeit verbunden zu sein. Die Theorie lässt einen positiven Zusammenhang vermuten (Stiglitz und Weiss, 1981). Hier können jedoch geeignete Vertragsgestaltungen durch das Kreditinstitut eine Risikoerhöhung durch den beschränkt haftenden Unternehmer verhindert haben (Bester, 1985 und 1987). Die Koeffizienten der Jahresdummies sind allesamt hoch signifikant und positiv, was bedeutet, dass 1992 ein Jahr mit besonders wenigen Kreditausfällen war. Auffällig ist dabei, dass die Koeffizienten der konjunkturellen Schwächejahre 1997 und 1998 signifikant größer sind als die der Vorjahre. Die konjunkturelle Lage hat also einen bedeutenden Effekt auf die Kreditausfallwahrscheinlichkeit.

Das höchste Ausfallrisiko weist der Sektor Maschinenbau auf, das geringste kommt dem Baugewerbe zu. Dabei ist anzumerken, dass es sich beim Baugewerbe um eine Branche handelt, die in den betrachteten Jahren von der Wiedervereinigung besonders profitieren konnte. Maschinenbau, Verarbeitendes Gewerbe und die sonstigen Unternehmen unterscheiden sich nicht signifikant voneinander.

Tabelle 1: Ergebnisse des Probitmodells

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	-0,8744	0,8333	0,294
ln(Umsatz) ²	0,0370	0,0361	0,306
Eigenkapitalquote	-2,8844	0,4011	0,000
Cash flow	-0,4113	0,1735	0,018
Anlagendeckungsgrad	-0,1253	0,0684	0,067
Beschr. Haftung	-0,0251	0,1428	0,861
1993	0,3484	0,1937	0,072
1994	0,6876	0,1829	0,000
1995	0,6720	0,1824	0,000
1996	0,7729	0,1822	0,000
1997	1,4553	0,2349	0,000
1998	1,7664	0,3188	0,000
Verarb. Gew.	-0,0594	0,1349	0,660
Baugewerbe	-0,5983	0,2018	0,003
Handel	-0,1343	0,1613	0,405
Sonstige	-0,0346	0,1417	0,807
Konstante	4,7183	4,7828	0,324
Wald-Tests auf gemeinsame Signifikanz			
Variablen	chi²	d.o.f	p-Wert
Umsatz	1,23	2	0,54
Jahresdummies	61,18	6	0,00
Branchendummies	9,80	4	0,04
ges. Schätzung	200,86	16	0,00

Die Ergebnisse in Tabelle 1 können zwar qualitativ als Effekte auf die Ausfallwahrscheinlichkeit interpretiert werden, der numerische Wert dieser Effekte lässt sich aus dieser Darstellung jedoch nicht ablesen. Während die geschätzten Koeffizienten im linearen Regressionsmodell nämlich unmittelbar als marginale Effekte bzw., bei logarithmierten Größen, in Elastizitäten zu interpretieren sind, müssen diese beim binären Probitmodell zunächst noch berechnet werden.

An dieser Stelle sei angemerkt, dass die absolute Größe selbst einer standardisierten – also eines auf einen Mittelwert von 0 und eine Standardabweichung von 1 normierten – Variable keineswegs Auskunft über die „Trennkraft“ einer Variablen gibt, wie es Backhaus et al. (2000, S. 113) formulieren. Sofern „Trennkraft“ die Präzision bezeichnet, mit der ein Koeffizient gemessen wird, dann ist die *t*-Statistik – also der Quotient von Koeffizient und dessen Standardabweichung – aufschlussreich, die angibt, ob ein Koeffizient überhaupt signifikant verschieden von 0 ist. Eine letzte Anmerkung zur Größe von Koeffizienten gilt den Koeffizienten von Dummyvariablen. Ein Absolutwert eines Dummyvariablen-Koeffizienten von über 3 (über 5,44) in Probit- (Logit-)modellen deutet auf eine

Fehlspezifikation des Modells weil – vereinfacht gesagt – die Auswahlwahrscheinlichkeit z. B. beim Probitmodell im Falle eines Wertes von 3 (-3) gleich 1 (0) ist, die übrigen Variablen der Schätzung gar keine Rolle mehr spielen.

Der Effekt einer einprozentigen Erhöhung der Variablen k im Vektor der erklärenden Variablen x_i ist – unter der Normierung $s=0$ und $\sigma=1$ – gegeben durch:

$$(3) \quad \begin{aligned} \frac{\partial P(\text{Ausfall}_i = 1 | x_i)}{\partial x_{ik}} &= f(-x_i\beta) \beta_k \\ \frac{\partial P(\text{Ausfall}_i = 0 | x_i)}{\partial x_{ik}} &= -f(-x_i\beta) \beta_k, \end{aligned}$$

wobei $f(\cdot)$ im Probitfall die Dichte der Standardnormalverteilung angibt und im Logitfall die der logistischen Verteilung. Gleichung (3) macht deutlich, dass der numerische Effekt auf die Wahrscheinlichkeitsveränderung vom Wert der Dichtefunktion $f(\cdot)$ an der Stelle $-x_i\beta$ abhängt. Dennoch lässt sich aus dem Vorzeichen des Koeffizienten β_k der Effekt auf die Wahrscheinlichkeitsveränderung eindeutig ablesen: Ein positives (negatives) Vorzeichen von β_k , bedeutet, dass die Variable x_k einen positiven (negativen) Effekt auf die Ausfallwahrscheinlichkeit hat. Ein Berechnen der marginalen Effekte macht natürlich nur für kontinuierliche Variablen Sinn. Möchte man hingegen den numerischen Effekt der Veränderung einer Dummy-Variablen analysieren, dann wird die geschätzte Ausfallwahrscheinlichkeit für den Wert der Dummy-Variablen mit dem Wert 1 mit der Ausfallwahrscheinlichkeit für den Wert der Dummy-Variablen mit dem Wert 0 verglichen:

$$(4) \quad P(\text{Ausfall}_i = 1 | x_{it} = 1, x_i) - P(\text{Ausfall}_i = 1 | x_{it} = 0, x_i) = F(-x_i\beta | x_{it} = 0) - F(-x_i\beta | x_{it} = 1),$$

wobei x_{it} eine Dummy-Variable bezeichnet. Ebenso wie für den Fall kontinuierlicher Variablen gibt das Vorzeichen des Koeffizienten β_1 Auskunft über den Effekt der erklärenden Dummy-Variable x_{it} . Gleichungen (3) und (4) implizieren, dass es für jedes Individuum i (bzw. für jeden Kreditnehmer i) einen marginalen Effekt gibt. In der Praxis werden die marginalen Effekte daher oft am Mittelwert der erklärenden Variablen berechnet.

Viele Standard-Softwareprogramme wie z. B. Limdep und STATA berechnen die marginalen Effekte standardmäßig sowohl für kontinuierliche als auch für diskrete Variablen. Tabelle 2 weist die marginalen Effekte des bereits in Tabelle 1 dargestellten Modells aus. Die in Tabelle 2 dargestellten Standardfehler wurden mit der „Delta“-Methode berechnet, die in Greene (1997, Kap. 6.7.5) näher beschrieben ist.

Anstatt der Koeffizienten werden in Tabelle 2 die marginalen Effekte dargestellt. Erhöht sich beispielsweise die Eigenkapitalquote um einen Prozentpunkt, so sinkt die Ausfallwahrscheinlichkeit um 0,87 Prozentpunkte.

Tabelle 2: Marginale Effekte der Probitschätzung

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	-0,2640	0,2515	0,294
ln(Umsatz) ²	0,0112	0,0109	0,306
Eigenkapitalquote	-0,8708	0,1186	0,000
Cash flow	-0,1242	0,0518	0,018
Anlagendeckungsgrad	-0,0378	0,0206	0,067
Beschr. Haftung	-0,0075	0,0425	0,861
1993	0,1135	0,0670	0,072
1994	0,2341	0,0666	0,000
1995	0,2281	0,0662	0,000
1996	0,2656	0,0669	0,000
1997	0,5306	0,0763	0,000
1998	0,6210	0,0798	0,000
Verarb. Gew.	-0,0178	0,0401	0,660
Baugewerbe	-0,1465	0,0379	0,003
Handel	-0,0391	0,0452	0,405
Sonstige	-0,0104	0,0421	0,807

Die marginalen Effekte des Cash Flow und der Anlagendeckung sind deutlich geringer. Erhöht sich etwa die Anlagendeckung um einen Prozentpunkt, so verringert sich die Ausfallwahrscheinlichkeit um 0,04 Prozentpunkte. Die Ausfallwahrscheinlichkeit eines Unternehmens des Baugewerbes ist um 14,9 Prozentpunkte geringer als die eines Unternehmens der Referenzbranche Maschinenbau.

Das binäre Probitmodell kann – ebenso wie das im nächsten Abschnitt beschriebene geordnete Probitmodell – Aussagen darüber treffen, wie hoch die Wahrscheinlichkeit ist, dass ein Kreditnehmer mit bestimmten Eigenschaften ausfällt. Will man z. B. wissen, wie hoch die Ausfallwahrscheinlichkeit eines Handelsunternehmens ohne Haftungsbeschränkung im Jahr 1995 ist, das 220 000 TDM Jahresumsatz aufweist, eine Eigenkapitalquote von 20 Prozent und einen dynamischen Cash Flow von 58 Prozent sowie eine Anlagendeckung von 80 Prozent hat, kann die Ausfallwahrscheinlichkeit anhand von Gleichung (2) wie folgt berechnet werden:

$$\begin{aligned}
 P(\text{Ausfall}_i = 1) &= 1 - F[-\{\ln(220\,000) \cdot \beta_{\ln(\text{Umsatz})} + \ln(220\,000)^2 \cdot \beta_{\ln(\text{Umsatz})^2} \\
 &+ 0,2 \cdot \beta_{\text{Eigenkapitalquote}} + 0,58 \cdot \beta_{\text{Cash Flow}} + 0,8 \cdot \beta_{\text{Anlagendeckung}} + 0 \cdot \beta_{\text{bhaft}} + \beta_{1995} + \beta_{\text{Handel}} + \\
 &\beta_{\text{Konstante}}\}] = 1 - F[-\{12,3 \cdot (-0,87) + 151,3 \cdot 0,04 + 0,2 \cdot (-2,88) + 0,58 \cdot (-0,41) + 0,8 \cdot (- \\
 &0,13) + 0,67 - 0,13 + 4,69\}] = 1 - F[0,3368] = 0,37
 \end{aligned}$$

Der oben beschriebene Kredit fällt also mit 37-prozentiger Wahrscheinlichkeit aus.

Trotz der Einfachheit des Probitmodells und seiner langjährigen Verwendung in den verschiedenen Teildisziplinen der Wirtschaftswissenschaften wird das Kreditausfallrisiko vielfach immer noch mit der Methode der kleinsten Quadrate (KQ) geschätzt. Der KQ-Schätzer sollte jedoch aus zweierlei Gründen nicht zur Modellierung von Ausfallwahrscheinlichkeiten verwendet werden. Zum einen erfüllt der Fehlerterm nicht mehr die Annahme eines identisch verteilten, also heteroskedastiefreien, Fehlerterms des linearen Regressionsmodells, was zu einer fehlerhaften Schätzung der Varianz-Kovarianzmatrix führt,⁸ zum anderen kann eine KQ-Schätzung zu logisch inkonsistenten Ergebnissen führen, da nicht gewährleistet ist, dass die geschätzten Ausfallwahrscheinlichkeiten zwischen 0 und 1 liegen. Hätten wir das in Tabelle 1 dargestellte Modell mit Hilfe kleinster Quadrate geschätzt, so hätte sich für einen Kreditnehmer eine Kreditausfallwahrscheinlichkeit größer als 1 und für 66 Kreditnehmer eine Kreditausfallwahrscheinlichkeit kleiner als 0 ergeben.

3.2 Mehrzustandsmodelle: Geordnete Probitmodelle

Wie eingangs erwähnt, sind binäre Logit- und Probitmodelle leicht erweiterbar, um mehrere Zustände abbilden zu können. Genau wie beim binären Probit wird beim geordneten Probitmodell davon ausgegangen, dass die Zustände, die ein Kredit aufweisen kann, von der Größe der unbeobachtbaren Variable $Ausfall_i^*$ abhängen. Kann ein Kredit drei Zustände annehmen, z. B. volle Rückzahlung ($Ausfall_i = 0$), teilweiser Ausfall ($Ausfall_i = 1$) und vollständiger Ausfall ($Ausfall_i = 2$), so wird beim geordneten Probitmodell davon ausgegangen, dass ein vollständiger Ausfall dann eintritt, wenn die unbeobachtbare Variable $Ausfall_i^*$ über einer oberen Schwelle s_2 liegt. Liegt der Wert der latenten Variablen unterhalb einer Schwelle s_1 , so wird der Kredit vollständig zurückgezahlt. Sofern die latente Variable zwischen den beiden Schwellen liegt, wird der Kredit teilweise zurückgezahlt:

$$(6) \quad \begin{aligned} & 0 \text{ (vollständiger Ausfall) wenn } Ausfall_i^* = x_i\beta + \varepsilon_i > s_2 \\ & Ausfall_i = \begin{cases} 1 \text{ (teilweiser Ausfall) wenn } s_1 \leq Ausfall_i^* < s_2 \\ 2 \text{ (kein Ausfall) wenn } Ausfall_i^* \leq s_1. \end{cases} \end{aligned}$$

Genau wie im binären Modell wird die latente Variable auch hier durch einen linearen Zusammenhang zwischen dem Vektor der erklärenden Variablen x_i sowie einem additiven, als identisch und unabhängig verteilten Störterm beschrieben. Ebenso müssen auch hier Restriktionen auferlegt werden. STATA wählt die

8 Mit anderen Worten: Jede statistische Inferenz ist fehlerhaft, es kann z. B. nicht mehr von „signifikanten“ oder „insignifikanten“ Koeffizienten gesprochen werden. Allerdings kann dieses Problem leicht mit der Verwendung „verallgemeinerter kleinster Quadrate“ behoben werden, da die Form der Heteroskedastie bekannt ist.

Restriktion $\sigma = 1$ und setzt den Wert der Konstanten auf 0 (es werden alle Schwellenwerte s mitgeschätzt), Limdep normiert σ ebenfalls auf 1 und setzt die untere Schwelle s_1 auf 0 (es wird der Koeffizient der Konstanten mitgeschätzt). Die Interpretation der geschätzten Parameter erfolgt analog zum binären Probitmodell. Ein positiver Koeffizient bedeutet, dass die korrespondierende erklärende Variable einen positiven Einfluss auf die unbeobachtbare Variable $Ausfall_i^*$ hat und somit die Wahrscheinlichkeit, dass ein „besserer“ Zustand erreicht wird, erhöht. Ebenso wie im binären Fall können beim geordneten Probitmodell leicht marginale Effekte berechnet werden. Dabei erhält man im Fall von drei Kategorien für jedes Individuum drei marginale Effekte:

$$(7) \quad \begin{aligned} \frac{\partial P(Ausfall_i = 2 | x_i)}{\partial x_{ik}} &= f(s_2 - x_i \beta) \beta_k \\ \frac{\partial P(Ausfall_i = 1 | x_i)}{\partial x_{ik}} &= (f(s_1 - x_i \beta) - f(s_2 - x_i \beta)) \beta_k, \\ \frac{\partial P(Ausfall_i = 0 | x_i)}{\partial x_{ik}} &= -f(s_1 - x_i \beta) \beta_k. \end{aligned}$$

Dabei wird deutlich, dass beim geordneten Probitmodell die Richtung der Wahrscheinlichkeitsveränderung bei einer Veränderung der erklärenden Variablen x_k nur bei den äußeren Kategorien durch das Vorzeichen des jeweiligen Koeffizienten bestimmt ist.

Tabelle 3 weist die Schätzergebnisse des geordneten Probitmodells für das oben beschriebene Modell aus. Dabei fällt auf, dass die Schwellenwerte s_1 und s_2 sehr unpräzise geschätzt sind und sich mithin nicht signifikant voneinander unterscheiden. Ursächlich dafür sind die Umsatzvariablen. Möglicherweise werden hier Umsatzgrößen-Gruppeneffekte auf die Schwellenwerte übertragen. Wird eine der beiden, ohnehin sowohl gemeinsam als auch getrennt insignifikanten, Umsatzvariablen weggelassen, so unterscheiden sich die beiden Schwellenwerte signifikant voneinander. Die übrigen Koeffizienten bleiben beinahe nahezu unverändert. In der Praxis und in der Wissenschaft würde man nun an dieser Stelle das Modell ohne die quadrierte Umsatzvariable ausweisen. Aus Gründen der Vergleichbarkeit mit den übrigen in diesem Beitrag ausgewiesenen Schätzergebnissen wird jedoch die Spezifikation mit beiden Umsatzvariablen besprochen und ausgewiesen.

Die Ergebnisse der Ordered-Probit-Schätzung unterscheiden sich qualitativ kaum von denen der einfachen Probitschätzung.⁹ Die Vorzeichen der Koeffizienten im geordneten Probitmodell geben die Richtung an, mit denen sich die Wahrscheinlichkeiten der Randkategorien verändern. In diesem Fall sind die

9 Dabei ist zu beachten, dass Ergebnisse unterschiedliche Logit- und Probitmodelle aufgrund der Skalierung durch die Standardabweichung der Normierung des Fehlerterms streng genommen nicht verglichen werden können.

Randkategorien die Wahrscheinlichkeit, mit der keine Probleme auftreten und die Wahrscheinlichkeit, mit der erhebliche Probleme auftreten. Über die dazwischen liegenden Kategorie, das heißt, die Wahrscheinlichkeit, mit der es zu einem Teilausfall kommt, kann anhand dieser Tabelle keine Auskunft gegeben werden.

Tabelle 3: Ergebnisse des geordneten Probitmodells

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	-0,7219	0,7946	0,364
ln(Umsatz) ²	0,0325	0,0345	0,345
Eigenkapitalquote	-2,7826	0,3840	0,000
Cash flow	-0,4819	0,1755	0,006
Anlagedeckungsgrad	-0,1195	0,0660	0,070
Beschr. Haftung	-0,0717	0,1355	0,597
1993	0,3579	0,1905	0,060
1994	0,5441	0,1810	0,003
1995	0,5853	0,1803	0,001
1996	0,7332	0,1784	0,000
1997	1,3938	0,2205	0,000
1998	1,7706	0,2912	0,000
Verarb. Gew.	-0,0234	0,1293	0,857
Baugewerbe	-0,5796	0,1963	0,003
Handel	-0,1505	0,1548	0,331
Sonstige	-0,0637	0,1349	0,637
s ₁	-3,5435	4,5621	0,781
s ₂	-2,9887	4,5618	0,744
Wald-Tests auf gemeinsame Signifikanz			
Variablen	chi²	d.o.f	p-Wert
Umsatz	1,16	2	0,56
Jahresdummies	67,01	6	0,00
Branchendummies	9,91	4	0,04
ges. Schätzung	205,49	16	0,00

Ein negatives Vorzeichen, wie im Fall der Eigenkapitalquote, bedeutet, dass mit einer Erhöhung dieser Variablen eine Verringerung der Wahrscheinlichkeit eines vollständigen Kreditausfalls verbunden ist. Gleichzeitig bedeutet das, dass sich die Wahrscheinlichkeit erhöht, mit der keine Probleme auftauchen. Aus der Betrachtung der marginalen Effekte lassen sich zusätzliche Informationen ziehen.

Im geschätzten Beispiel vergrößert sich bei einer Erhöhung der Eigenkapitalquote um einen Prozentpunkt die Wahrscheinlichkeit, dass kein Problem auftaucht, um 0,80 Prozentpunkte. Die Wahrscheinlichkeit, dass ein Teilausfall stattfindet, verringert sich um 0,36 Prozentpunkte und die Wahrscheinlichkeit,

mit welcher der Kredit vollständig ausfällt, verringert sich um 0,44 Prozentpunkte. Die Wahrscheinlichkeitsveränderungen (+0,80, -0,36, -0,44) ergänzen sich dabei per Konstruktion zu 0, da die Fläche unter einer Verteilungsfunktion in der Summe immer unverändert bleiben muss. Sie beträgt als grundlegende Eigenschaft von Wahrscheinlichkeitsverteilungen in der Summe 1. Genauso wie im binären Probitmodell sind die marginalen Effekte der Variablen Anlagendeckung und dynamischer Cash-Flow deutlich geringer als die der Eigenkapitalquote.

Tabelle 4: Marginale Effekte stetiger Variablen im geordneten Probitmodell

	kein Ausfall	teilw. Ausfall	vollst. Ausfall
ln(Umsatz)	0,2081	-0,0940	-0,1141
ln(Umsatz) ²	0,0094	0,0042	0,0051
Eigenkapitalquote	0,8021	-0,3624	-0,4396
Cash flow	0,1389	0,0628	-0,0761
Anlagendeckungsgrad	0,0344	0,0156	-0,0189

Auch für Dummy-Variablen lassen sich – wie auch schon im Fall des binären Probitmodells – marginale Effekte berechnen. Für den Einfluss einer Haftungsbeschränkung ergeben sich folgende Zahlen. Eine Haftungsbeschränkung verringert die Wahrscheinlichkeit, mit der keine Zahlungsschwierigkeiten auftreten um 0,02 Prozentpunkte, die Wahrscheinlichkeit eines Auftretens von leichten Problemen wird um 0,01 Prozentpunkte erhöht und die Wahrscheinlichkeit für erhebliche Vertragsstörungen steigt um 0,01 Prozentpunkte.

An dieser Stelle wird der Vorteil eines geordneten Probit-Modells gegenüber dem binären Probitmodell deutlich. Der Kreditbeurteiler erhält mehr Information über die Wahrscheinlichkeit des Eintretens unterschiedlich schwerer Probleme. Da die Schwellen zwischen unterschiedlich schweren Problemkategorien geschätzt werden, könnte so ein Rating-System (ein Scoring-System) konstruiert werden, dessen Rating-Klassen bestimmten Problemkategorien entsprechen und damit deutlich aussagekräftiger wären.¹⁰

3.3 Spezifikationstests

Einer der wichtigsten Gesichtspunkte bei der Kreditrisikomessung ist zweifellos die Frage, ob die gewählte Spezifikation korrekt ist – also die Annahmen des Modells erfüllt sind – und wie gut die Spezifikation einen Kreditausfall vorher-sagen kann.

¹⁰ Dies ist natürlich nur dann der Fall, wenn die einzelnen Problemklassen auch eindeutig identifiziert werden können und sich die Schwellenwerte signifikant voneinander unterscheiden. Um ein Rating-System zu konzipieren, würde im vorliegenden Beispiel also der quadrierte Umsatzterm in der Schätzung ausgelassen.

Bevor wir zu Spezifikationstests und Gütemassen kommen, möchten wir noch einige allgemeine Anmerkungen zum Gebrauch des R^2 als Gütemass einer Schätzung machen. Diese Anmerkungen gelten in weiten Teilen sowohl für das lineare Regressionsmodell als auch für Maximum-Likelihood-Schätzer. In vielen empirischen Arbeiten findet die Größe des R^2 eine hohe Aufmerksamkeit. Im linearen Regressionsmodell wird das R^2 dabei häufig als Maß für die Qualität des *statistischen* Modells interpretiert, während es in Wahrheit nichts anderes darstellt als ein Maß für die Güte der linearen Approximation. Bei Maximum-Likelihood-Schätzungen ist das sogenannten Pseudo- R^2 lediglich ein Maß für den Informationsgehalt, der in den erklärenden Variablen steckt und sagt über die Qualität eines Modells gar nichts aus. Tatsächlich ist ein sehr hoher Wert des Pseudo- R^2 in vielen Fällen ein Anzeichen von Fehlspezifikation des Modells. Aus mehreren weiteren Gründen ist die hohe Bedeutung, der dem R^2 in vielen betriebswirtschaftlichen Arbeiten zukommt, ungerechtfertigt: (i) Das R^2 geht davon aus, dass die gewählte Spezifikation korrekt ist. So wiesen viele zeitreihenökonomische Arbeiten der 70er und Anfang der 80er Jahre häufig ein R^2 von 0,99 aus, z. B. der wichtige Beitrag von Nadiri und Rosen (1969) zu dynamischen Faktornachfragesystemen. Letztlich handelte es sich jedoch meistens lediglich um „spurious regressions“, um Scheinregressionen, wie Granger und Newbold (1974) solche Regressionen nannten, bei denen sowohl die zu erklärende als auch die erklärenden Variablen einem gemeinsamen Zeittrend folgten. (ii) Das R^2 ist nicht invariant gegenüber affinen Transformationen der zu erklärenden Variable. Obwohl sich bei linearen Transformation der zu erklärenden Variablen identische Koeffizienten der erklärenden Variablen ergeben, unterscheidet sich das R^2 des nicht transformierten Modells von dem des transformierten Ansatzes. (iii) Das R^2 und auch das um die Anzahl der Regressoren korrigierte R^2 , ist leicht manipulierbar.¹¹

Aufgrund der Tatsache, dass das R^2 im linearen Regressionsmodell den Anteil der durch die lineare Approximation erklärten Varianz an der Gesamtvarianz der zu erklärenden Variable misst, macht es auch wenig Sinn, Schwellenwerte zu benennen, ab denen ein Modell als „gut“ bezeichnet werden kann. So liegt das R^2 linearer Regressionen mit Zeitreihendaten fast immer wesentlich höher als das R^2 linearer Regressionen mit Individualdaten (z. B. Unternehmensdaten, Personendaten, Daten aus Kreditakten). Dies liegt ganz einfach daran, dass Individualdaten sehr viel stärker streuen – also eine höhere Varianz aufweisen – als Zeitreihendaten.¹²

Doch selbst wenn das R^2 nicht mit den angesprochenen Problemen behaftet wäre, erscheint es ratsam, zunächst auf Möglichkeiten, die Robustheit der Modellspezifikation zu testen, einzugehen. Denn erst wenn die Robustheit des

11 Leamer (1978, Kap. 3) nimmt dazu ausführlich Stellung.

12 Ausnahmen sind z. B. extrem volatile Finanzmarktdaten, z. B. Intratagesdaten (s. die Sonderausgabe des Journal of Business and Economic Statistics, Band 18 Nr. 2, April 2000).

Schätzmodells gewährleistet ist, können Gütemaße überhaupt sinnvoll angewendet werden.

Während Heteroskedastie im linearen Regressionsmodell die Konsistenz der Koeffizienten unbeeinflusst lässt, werden die geschätzten Koeffizienten bei binären und geordneten Probit- bzw. Logitmodellen inkonsistent.¹³ Die geschätzten Koeffizienten der erklärenden Variablen sind also „falsch“. Heteroskedastie hat weitaus stärkere Auswirkungen auf Logit- und Probitmodelle als auf lineare Regressionsmodelle. Dies liegt daran, dass, wie aus Gleichung (2) ersichtlich ist, nicht der Koeffizientenvektor β geschätzt wird, sondern das Verhältnis β/σ . Durch die Normierung $\sigma=1$ wird der Koeffizientenvektor zwar identifiziert, doch handelt es sich bei dieser Normierung lediglich um eine identifizierende *Restriktion*. Liegt Heteroskedastie vor, z. B. der generellen Form $\sigma_i=h(\gamma z_i)$, wobei γ einen Koeffizientenvektor bezeichnet und z_i diejenigen Variablen bezeichnet, die Heteroskedastie hervorrufen, so wird anstatt des Parametervektors β das Verhältnis $\beta/\exp(z_i \gamma)$ geschätzt.¹⁴ Die Werte der Parametervektors variieren also mit der *iten* Beobachtung und mit den Werten der Variablen, die die Heteroskedastie verursachen. Ebenso führt eine Verletzung der Normalverteilungsannahme (bzw. der logistischen Verteilung beim Logit-Modell) zu Inkonsistenz der Schätzergebnisse, weil die Ausfallwahrscheinlichkeit über die Verteilungsannahme explizit modelliert wird.

Chesher und Irish (1987) schlagen auf „generalisierten Residuen“ aufbauende Tests auf Heteroskedastie und Nicht-Normalität der Residuen vor.¹⁵ Weil der Wert der latenten Variablen, in diesem Fall der Wert von Ausfall_i unbeobachtbar ist – also Residuen nicht direkt berechnet werden können –, können Residuentests wie sie für das lineare Regressionsmodell bestehen, nicht angewendet werden. Aus diesem Grund entwickeln Chesher und Irish (1987) generalisierte Residuen, die im binären Logit- bzw. Probitmodell folgende Form annehmen:

$$(8) \quad \varepsilon_i^g = \frac{\text{Ausfall}_i - F(x_i \beta)}{F(x_i \beta)(1 - F(x_i \beta))} f(x_i \beta).$$

Für das geordnete Logit- bzw. Probitmodell lautet die Formel wie folgt:

$$(9) \quad \zeta_i^g = \frac{f(s_{j-1} - x_i \beta) - f(s_j - x_i \beta)}{F(s_j - x_i \beta) - F(s_{j-1} - x_i \beta)},$$

wobei der Index j den j ten Schwellenwert angibt.

13 „Konsistenz“ bedeutet, dass die Präzision mit der die Parameter eines Modells geschätzt werden mit zunehmender Stichprobengröße genauer wird.

14 Laisney et al. (1991) demonstrieren, zu welchen starken Abweichungen es zwischen den Koeffizientenvektoren β eines homoskedastischen und eines heteroskedastischen Logitmodells kommen kann.

15 Vielfach werden die generalisierten Residuen im Sinne von Chesher und Irish (1987) mit den simulierten Residuen (Gourieroux et al., 1987) verwechselt. Ein grundlegender Unterschied zwischen beiden Ansätzen ist, dass die simulierten Residuen durch Zufallsziehungen aus gestutzten Verteilungen erzeugt werden während die generalisierten Residuen direkt berechnet werden.

Sowohl der Test auf Heteroskedastie als auch der Test auf Normalität können über eine Hilfsregression durchgeführt werden. Die Nullhypothese lautet $\gamma=0$, es liegt also keine Heteroskedastie vor. Die korrespondierende Teststatistik für einen Test auf Heteroskedastie ergibt sich als Anzahl der Beobachtungen der Logit- bzw. Probitschätzung, N , multipliziert mit dem nicht korrigierten R^2 einer linearen Regression von (i) den Interaktionen zwischen dem generalisierten Residuum und den erklärenden Variablen des Logit- bzw. Probitmodells, $\varepsilon_i^s x_i$, und (ii) der Interaktion zwischen dem generalisierten Residuum und sowohl dem gefitteten Wert der Ursprungsschätzung, $x_i \beta$, als auch mit den Variablen, die im Verdacht stehen, Heteroskedastizität hervorzurufen, $(\varepsilon_i^s x_i \beta) z_i$, auf einen Vektor von Einsen ohne den Einschluss einer Konstanten.¹⁶ Analoges gilt für die generalisierten Residuen der geordneten Modelle. Die sich als $N \cdot R^2$ ergebende Teststatistik ist χ^2 verteilt mit Anzahl der Freiheitsgraden gleich Anzahl der im Verdacht Heteroskedastie hervorzurufen stehenden Variablen z .¹⁷

Wie beim Breusch-Pagan Test im linearen Regressionsmodell wird in der Praxis häufig davon ausgegangen, dass diejenigen Variablen, die die Wahrscheinlichkeitsauswahl bestimmen, auch diejenigen sind, die Heteroskedastie hervorrufen können. Wir gehen also davon aus, dass alle Variablen x , die den Kreditausfall erklären, Heteroskedastie verursachen könnten: $x=z$. Entsprechend regressieren wir $\varepsilon_i^s x_i$ und $(\varepsilon_i^s x_i \beta) x_i$ auf einen Einservektor. Das sich aus dieser Hilfsregression ergebende R^2 lautet für das binäre Probitmodell 0,0223. Die Teststatistik lautet entsprechend $N \cdot R^2 = 944 \cdot 0,0223 = 21,0395$. Die kritischen Werte einer χ^2 -Verteilung mit 16 Freiheitsgraden sind 23,54 auf dem 10-Prozent, 26,3 auf dem 5-Prozent und 32 auf dem 1-Prozent Signifikanzniveau. Die Teststatistik für das binäre Probitmodell von Abschnitt 3 ist also signifikant kleiner als die kritischen Werte. Homoskedastizität kann also nicht verworfen werden. Das marginale Signifikanzniveau beträgt 0,1770. Auch für das geordnete Probitmodell kann Homoskedastizität auf den konventionellen Signifikanzniveaus nicht verworfen werden.

Wenn Homoskedastizität verworfen werden muss, dann kann beobachtbare Heteroskedastie explizit mitgeschätzt werden. Anstelle von

$$(2) \quad P(\text{Ausfall}_i = 1 | x_i) = 1 - F\left(\frac{\varepsilon_i \leq s - x_i \beta}{\sigma}\right)$$

16 Es handelt sich hierbei um eine sogenannte „künstliche Regression“, Davidson und Mackinnon (1993, Kap. 6) gehen näher auf dieses in der Ökonometrie sehr wichtige Verfahren ein. Die Terme $\varepsilon_i^s x_i$ bzw. $(\varepsilon_i^s x_i \beta) z_i$ bezeichnen die Scores (1. Ableitungen) der Log-Likelihoodfunktion nach β bzw. γ .

17 Aufgrund der Tatsache, dass Logit- und Probitmodelle meist auf Individual- und nicht auf Zeitreihendaten angewendet werden, spielt Autokorrelation bei Auswahlmodellen eine geringere Rolle als bei Zeitreihenmodellen. Gourieroux et al. (1985) schlagen jedoch einen Test auf Autokorrelation bei binären und geordneten Auswahlmodellen vor.

würde dann

$$(2') \quad P(\text{Ausfall}_i = 1 | x_i) = 1 - F\left(\frac{\varepsilon_i \leq s - x_i \beta}{h(z_i \gamma)}\right)$$

geschätzt. Standardsoftwareprogramme wie STATA und Limdep spezifizieren die Funktion $h(\cdot)$ als Exponentialfunktion. Analoges gilt für die geordneten Modelle.¹⁸ Ein Test auf $\gamma=0$ entspricht ebenfalls einem Test auf Homoskedastizität.

Der Test auf Normalität der Residuen wird ebenfalls über eine Hilfsregression durchgeführt. In diesem Falle werden die Interaktionen $\varepsilon_i^s x_i$, $\varepsilon_i^s (x_i \beta)^2$ und $(\varepsilon_i^s x_i \beta)^3$ auf einen Einservektor regressiert. Die Koeffizienten der letzten beiden Terme entsprechen Schiefe und Wölbung der Verteilung der Residuen. Die Teststatistik lautet $N \cdot R^2$ aus der Hilfsregression, sie ist χ^2 verteilt mit zwei Freiheitsgraden. Das R^2 aus der Hilfsregression für das binäre Probitmodell lautet 0,0006, die Teststatistik beträgt also 0,5664. Die entsprechenden kritischen Werte der χ^2 -Verteilung lauten 4,61, 5,99 und 9,21 auf dem 10, 5 und 1 Prozent Signifikanzniveau. Normalität kann für das oben spezifizierte Modell also nicht verworfen werden. Normalität kann auch für das geordnete Probitmodell nicht verworfen werden. Es liegen, außer dem bereits besprochenen Problem der sich nicht signifikant unterscheidenden Schwellenwerten im geordneten Probitmodell, keine Anzeichen für Fehlspezifikationen vor.

Eine ähnlich einfache Lösung wie für das Vorliegen von Heteroskedastie gibt es beim Vorliegen von Nichtnormalität nicht. Gabler et al. (1993) stellen einen semiparametrischen Schätzer der Ausfallwahrscheinlichkeiten vor. Dabei wird, grob gesagt, die Funktion $F(\cdot)$ aus Gleichung (2) nichtparametrisch, also ohne die Spezifikation von Parametern auf der Grundlage der empirischen Verteilung der Residuen, geschätzt. Den Daten wird also nicht eine funktionale Form wie z. B. Normalverteilung auferlegt, vielmehr bestimmen die Daten selbst über den korrekten funktionalen Zusammenhang.

Eine ergänzende Referenz zu Tests auf Verteilungsannahmen bieten Chesher et al. (1985). Tests auf Fehlspezifikationen in binären und geordneten Wahrscheinlichkeitsmodellen, die auf Tests der Informationsmatrix beruhen, werden von Laisney et al. (1991) angewendet.¹⁹ Informationsmatrixtests auf Normalität im geordneten Wahrscheinlichkeitsmodellen werden von Glewwe (1997) und Weiss (1997) vorgeschlagen. Einen Überblick über Informationsmatrixtests bieten Gourieroux und Montfort (1995, Kap. 18.5).

18 Eine Anwendung dieser Spezifikation eines geordneten Probitmodells mit Heteroskedastie für Finanzmarktdaten findet sich in Hausman et al. (1992) sowie Kaiser (1997).

19 Die Informationsmatrix ist definiert als negative des Erwartungswerts der Hesse-Matrix (der Matrix der zweiten Ableitungen der Log-Likelihoodfunktion). Sie ist ein Schätzer der Varianz-Kovarianzmatrix des Koeffizientenvektors.

3.4 Gütemaße

Da es bei den binären und geordneten Wahrscheinlichkeitsmodellen aufgrund der Unbeobachtbarkeit des Fehlerterms kein echtes Analogon zum R^2 des linearen Regressionsmodells gibt, werden zur Überprüfung der Schätzgüte häufig sogenannte „prediction/realization tables“ verwendet. Dabei wird untersucht, wie oft – in unserem Fall – ein Kreditausfall aus dem geschätzten Modell heraus korrekt prognostiziert wurde.

Beim binären Probitmodell wird Zustand 1 dann vorhergesagt, wenn der Wert der linearen Vorhersage (der Wert $x_i \hat{\beta}$) den Wert 0 übersteigt und umgekehrt:

$$(10) \quad \begin{aligned} \hat{Ausfall}_i^* &= 1 \text{ wenn } x_i \hat{\beta} > 0 \\ \hat{Ausfall}_i^* &= 0 \text{ wenn } x_i \hat{\beta} \leq 0, \end{aligned}$$

wobei $\hat{Ausfall}_i^*$ und $x_i \hat{\beta}$ geschätzte Werte bezeichnen.

Problematisch an diesem Ansatz ist, dass sowohl das binäre als auch das geordnete Probitmodell per Konstruktion immer denjenigen Zustand am besten beschreiben, der am häufigsten eintritt. Tritt ein Zustand sehr selten ein, so wird dieser weniger präzise vorausgesagt. Auch aus diesem Grund sollte jede der Auswahlkategorien (der Risikozustände eines Kredites) mit mindestens fünf Prozent der Beobachtungen besetzt sein.²⁰ Wenn ein Zustand seltener als in fünf Prozent der Fälle eintritt, so müssen einzelne Zustände entweder zusammengefasst oder ggf. sogenannte „rare event“-Modelle verwendet werden (z. B. King und Zeng, 1999).

Ein weiteres Problem dieser prediction/realization Tabellen liegt darin, dass sie Vorhersagen, die z. B. auf einem Wahrscheinlichkeitswert von 0,51 basieren, die gleiche Bedeutung zumessen wie solchen, die auf einem Wahrscheinlichkeitswert von 0,99 beruhen. Diese Vorhersagen würden beide zu den Ausfällen gezählt werden, obwohl der Ausfall mit deutlich unterschiedlicher Wahrscheinlichkeit eintritt.²¹

Zu den in Tabelle 1 abgebildeten Ergebnissen korrespondiert der „prediction/realization table“ in Tabelle 5. Die Trefferquote bei Nicht-Problemfällen beträgt somit 94 Prozent, bei Problemfälle hingegen nur 30 Prozent. Das bedeutet, dass 70 Prozent der Beobachtungen, in denen Probleme sichtbar wurden, fälschlicherweise als potenziell unproblematisch klassifiziert wurden. Diese Fehlerquoten spielen verständlicherweise in der Praxis eine erhebliche Rolle, da sie mit Kosten verbunden sind. Ein fälschlicherweise an einen schlechten Kunden verbogener Kredit kann ausfallen, ein fälschlicherweise abgelehnter guter Kredit-

²⁰ Dies entspricht, formal gesprochen, einer schlechten Identifikation der jeweiligen Ausfallwahrscheinlichkeiten (s. Blundell et al., 1993).

²¹ Veall und Zimmermann (1992) nehmen zur Verwendung von prediction/realization-Tabellen ausführlich Stellung.

kunde ist mit entgangenen Gewinnen verbunden. Verfügt die Bank über Schätzungen dieser Kosten bzw. der entgangenen Gewinne, so kann ihre auf der Scoring-Funktion aufbauende Entscheidungsregel kostenoptimal gestaltet werden.²²

Tabelle 5: Trefferquoten des binären Probitmodells

		vorhergesagt		
		Ausfall: ja	Ausfall: nein	
tatsächlich	Ausfall: ja	645 94,02%	41 5,98%	686
	Ausfall: nein	181 70,16%	77 29,84%	258
		826	118	944

Andere populäre Gütemaße beruhen auf Vergleichen der Werte der Log-Likelihoodfunktion des vollständig parametrisierten Modells, mit Werten der Log-Likelihoodfunktion eines Modells, das ausschließlich aus einer Konstanten und, beim geordneten Probitmodell, aus einer oder mehr Schwellenwerten besteht. So geben Standardsoftwareprogramme häufig das Bestimmtheitsmaß von McFadden (1974) an, dass sich wie folgt berechnet:

$$(11) \quad R_{MF}^2 = 1 - \frac{\ln(L^u)}{\ln(L^r)}$$

wobei $\ln(L^u)$ die Log-Likelihoodfunktion des unrestringierten Modells – des Modells mit allen erklärenden Variablen – und $\ln(L^r)$ die Log-Likelihoodfunktion des restringierten Modells – des Modells nur mit einer Konstanten – angibt. Per Konstruktion liegt dieses Maß, ebenso wie das klassische R^2 des linearen Regressionsmodells, zwischen 0 und 1. Allerdings deuten sehr hohe Werte dieses Bestimmtheitsmaßes auf eine Fehlspezifikation hin, weil es genau dann den Wert 1 annimmt, wenn der Maximum-Likelihood-Schätzer nicht existiert. Insofern überrascht es ein wenig, dass Backhaus et al. (2000, S. 116) unter Berufung auf Urban (1993, S. 62) davon sprechen, dass „bereits bei Werten von 0,2-0,4 von einer guten Modellanpassung gesprochen werden kann“. Im Umkehrschluss bedeutet dies, dass Werte von unter 0,2 auf eine schlechte Modellanpassung hindeuten. Das McFadden Pseudo- R^2 kann jedoch gar keine Aussage über eine Modellanpassung treffen. Es macht lediglich eine Aussage darüber, wie deutlich sich die Werte der Log-Likelihoodfunktionen voneinander unterscheiden. Es macht somit eine Aussage über den Informationsgehalt, der in den erklärenden Variablen steckt. Zudem machen Vergleiche mit dem R^2 aus dem linearen Regressionsmodell wenig Sinn, da es dort den Anteil der durch das Modell erklärten Varianz

²² Vgl. zu dieser Thematik Fahrmeier et al. (1984).

bestimmt. McFaddens Pseudo-R² und das R² der linearen Regression haben nichts miteinander gemein. Vor diesem Hintergrund ist es auch unverständlich, dass Eckey et al. (1995, S. 179) McFaddens Pseudo-R² als „Unsicherheitsmaß“ bezeichnen. Aufgrund einer impliziten Bestrafung hoher Stichprobengrößen liegt das Pseudo-R² bei kleinen Stichproben generell höher als bei großen Stichproben.

Für das binäre Probitmodell aus Tabellen 1 ergibt sich ein McFadden Pseudo-R² von 0,1814, für das geordnete Probitmodell aus Tabelle 3 eines von 0,1461. Aufgrund des größeren Informationsgehaltes des geordneten Probitmodells sollte man erwarten, dass das Pseudo-R² des geordneten Probitmodells erheblich größer sein sollte als das des binären Probitmodells. Dass dies hier nicht der Fall ist, weckt nicht gerade Vertrauen in das McFaddens Pseudo-R². Dabei sollte allerdings nicht vergessen werden, dass sich in unserem Beispiel zum Kreditrisiko die Schwellenwerte beim geordneten Probitmodell nicht signifikant unterscheiden, der Informationsgehalt der zusätzliche Kategorie also sehr gering ist.

Veall und Zimmermann (1992) diskutieren und testen verschiedene Gütemaße für das binäre Probitmodell, die ohne Einschränkung auf geordnete Probitmodelle übertragen werden können. Auf der Grundlage von Simulationsstudien kommen sie zu dem Ergebnis, dass das Gütemaß von McKelvey und Zavoina (1975) deutlich besser abschneidet als das Maß von McFadden und weitere Maße, die auf der grundlegenden Idee von McFadden beruhen. Das Bestimmtheitsmaß von McKelvey und Zavoina (1975) berechnet sich als:

$$(12) \quad R_{MZ}^2 = \frac{\sum_{i=1}^N \left(\hat{Ausfall}_i^* - \overline{\hat{Ausfall}_i^*} \right)^2}{\sum_{i=1}^N \left(\hat{Ausfall}_i^* - \overline{\hat{Ausfall}_i^*} \right)^2 - N}$$

Das Gütemaß nach McKelvey und Zavoina (1975) beträgt 0,6149 für das binäre Probitmodell und 0,9571 für das geordnete Probitmodell.

Veall und Zimmermann (1992) zeigen, dass dieses Bestimmtheitsmaß das unbeobachtbare Modell korrekt reproduziert, während die auf dem Likelihood-Quotienten-Prinzip basierenden Gütemaße die wahren Werte unterschätzen. Im Gegensatz zum pseudo R² von McFadden zeigt das Maß von McKelvey und Zavoina, dass die Spezifikation in Tabelle 3 der Spezifikation in Tabelle 1 überlegen ist, da die Berücksichtigung von drei Kategorien im geordneten Probitmodell mehr Information über die Verteilung der latenten Variable Ausfall zur Verfügung stellen als es das binäre Probitmodell mit nur zwei Kategorien vermag.

In einer aktuellen Arbeit schlägt Estrella (1998) ein neues Gütemaß, das ebenfalls im Likelihood-Quotienten-Prinzip seinen Ursprung hat, vor. Mit Hilfe von

Simulationen zeigt er, dass es anderen Gütemaßen in seiner Genauigkeit überlegen ist.²³ Dieses Gütemaß berechnet sich als

$$(13) \quad R_E^2 = 1 - \left(\frac{\ln(L'')}{\ln(L')} \right)^{-(2/N)\ln(L')}$$

und beträgt 0,2093 für das binäre und 0,2097 für das geordnete Probitmodell.

Aufgrund der Schwierigkeit, das Pseudo- R^2 zu interpretieren, bietet es sich an, die gemeinsame Signifikanz der erklärenden Variablen im binären oder geordneten Wahrscheinlichkeitsmodell zu testen. Ein einfacher und in Standardsoftwareprogrammen wie STATA und Limdep routinemäßig ausgewiesener Test auf gemeinsame Signifikanz ist der Likelihood-Ratio-Test (LR-Test), der sich als $-2 \cdot [\ln(L') - \ln(L'')]$ berechnet. Das Pseudo- R^2 ist also lediglich eine andere, jedoch schwieriger zu interpretierende, Darstellung dieses Tests auf Signifikanz der erklärenden Variablen. Diese LR-Teststatistik ist χ^2 verteilt mit Anzahl an Freiheitsgraden gleich der Differenz der Koeffizienten des restringierten und des unrestringierten Modells. Sie lautet für das binäre Probitmodell beispielsweise 200,86 für das binäre und 205,49 für das geordnete Probitmodell. Die korrespondierenden kritischen Werte der χ^2 -Verteilung auf dem 10, 5 und 1 Prozent Signifikanzniveau lauten 5,81, 7,96 und 9,31. Beide Koeffizientenvektoren sind also hoch signifikant verschieden von 0.

3.5 Erweiterungen

3.5.1 Paneldatenmodell

Die in den Abschnitten 2.1 und 2.2 dargestellten Modelle lassen sich in vielerlei Hinsicht erweitern. Eine nahe liegende und in vielen Softwareprogrammen, z. B. STATA und Limdep, bereits implementierte Erweiterung sind Logit- und Probitmodelle für Paneldaten. Als Paneldaten werden Daten bezeichnet, die sowohl eine Querschnitts- als auch eine Zeitreihendimension besitzen. Im Falle der Kreditrisikomessung handelt es sich dann um einen Paneldatensatz, wenn zu einem einzelnen Kredit i zu mehreren Zeitpunkten t Informationen vorliegen. In unserem Fall wurde – um relativ regelmäßig über den Zeitraum verteilte Beobachtungen pro Kredit zu erhalten – alle Ereignisse des Jahres zum Ende des Jahres zusammengefasst und als eine Beobachtung dargestellt.

Wesentliche Vorteile von Paneldatensätzen sind, dass sie (i) aufgrund der Querschnitts- und der Zeitreihendimension i.d.R. wesentlich mehr Datenpunkte umfassen als Querschnitts- oder Zeitreihendaten alleine und damit die Genauigkeit der Schätzung erhöhen, (ii) es ermöglichen, Fragestellungen anzugehen, die mit Zeitreihen- oder Querschnittsdaten alleine nicht zu beantworten sind. Bei-

²³ Estrella (1998) berücksichtigt dabei allerdings nicht das pseudo R^2 von McKelvey und Zavoina (1975).

spielsweise erlauben nur Paneldaten es, Aspekte der Geschichte eines Individuums zu beobachten. Ein letzter zentraler Vorteil von Paneldaten ist die Berücksichtigung von unbeobachtbarer Heterogenität, im Sinne von für den Analytisten unbeobachtbaren Unterschieden zwischen den Individuen. Paneldaten erlauben es, z. B. den Fehlerterm in eine individuenspezifische Komponente (unbeobachtbare Heterogenität) und eine zufällige Komponente zu zerlegen.

Das Analogon zu Gleichung (1) lautet für Paneldaten:

$$(1') \quad \text{Ausfall}_i = \begin{cases} 1 & \text{wenn } \text{Ausfall}_{it}^* = x_{it}\beta + \varepsilon_{it} = x_{it}\beta + \alpha_i + u_{it} \geq s \\ 0 & \text{sonst.} \end{cases},$$

wobei der Vektor der erklärenden Variablen x_{it} keine Konstante enthält und α_i einen individuenspezifischen Effekt repräsentiert. Dieser individuenspezifische Effekt wird als über die Zeit konstant angenommen und häufig als „unbeobachtbare Heterogenität“ bezeichnet. Wenn α_i über alle Individuen i konstant ist und Unabhängigkeit der Fehlerterme u_{it} zu verschiedenen Zeitpunkten besteht sowie beide Komponenten logistisch verteilt sind, dann ergibt sich das gewöhnliche binäre Logitmodell, das beim Vorliegen von Paneldaten als „gepoolte Schätzung“ bezeichnet wird.

Zwei prinzipielle Möglichkeiten zur Schätzung eines wie in Gleichung (1') dargestellten Modells existieren. Der erste Ansatz, das Fixed-Effects-Modell, geht von einer individuenspezifischen Konstanten α_i im Schätzmodell und einer Unabhängigkeit der Fehlertermkomponenten u_{it} aus. Natürlich kann aufgrund des damit, gegenüber dem gepoolten Modell bzw. dem Modell mit der allgemeinen Konstanten α , verbundenen Verlusts an Freiheitsgraden ein Modell mit N Individualeffekten nicht geschätzt werden. Chamberlain (1980) hat daher ein Verfahren vorgeschlagen, mit dem die unbeobachtbare Heterogenität auf andere Weise – durch Maximierung der konditionalen (im gegenwärtigen Fall konditional auf die Anzahl der ausgefallenen Kredite) anstelle der nicht konditionalen Likelihoodfunktion – entfernt werden kann. Daraus ergibt sich, dass in die fixed-effects Schätzung lediglich Beobachtungen aufgenommen werden können, die mindestens einmal ihren Status wechseln. Im uns vorliegenden Datensatz gibt es jedoch 383 Kredite oder 563 Beobachtungen (383 Kredite beobachtet zu verschiedenen Zeitpunkten), für die dies nicht der Fall ist. Bei einer Fixed-Effects-Schätzung würde also mehr als die Hälfte der Beobachtungen verloren gehen. Weil aber diese Kredite, die ihren Status nicht verändern, gerade für die Kreditrisikomesung von hoher Bedeutung sind, verzichten wir auf die weitere Darstellung des Fixed-Effects-Schätzers.²⁴

24 Für das Probitmodell existiert in Standardsoftwarepaketen bislang nur der Fixed-Effects-Schätzansatz während für das Logitmodell sowohl Fixed- als auch Random-Effects-Schätzer vorhanden sind, so dass sich die Diskussion von binären Paneldatenmodellen in diesem Abschnitt auf Logitmodelle konzentriert.

Der zweite Ansatz, das Random-Effects-Modell, geht davon aus, dass beide Komponenten, α_i und u_{it} mit Mittelwert 0 und unabhängig voneinander logistisch verteilt sind. Es gilt also für die Varianz des Fehlerterms $Var[\varepsilon_{it}] = 1 + \sigma_u^2$ und für die Korrelation der beiden Fehlertermkomponenten $Corr[\varepsilon_{it}, \varepsilon_{is}] = \sigma_u^2 / (1 + \sigma_u^2)$. Dabei handelt es sich bei σ_u^2 um einen zu schätzenden Parameter. Ein Likelihood-Ratio-Test auf σ_u^2 entspricht dabei einem Test auf Existenz von Random Effects.²⁵

Es ist a priori schwer zu sagen, ob für die konkrete empirische Anwendung das Fixed-Effects-Modell dem random effects Ansatz vorzuziehen ist. Der wesentliche Unterschied zwischen beiden Modellen liegt in den unterschiedlichen Annahmen bezüglich der Korrelation der zeitinvarianten Individualeffekte (den α_i 's) und den Residuen (den u_{it} 's) bzw. den erklärenden Variablen und den Individual-effekten. In der Praxis wird der Fixed-Effects-Schätzer häufig als überzeugender angesehen, was daran liegt, dass es vermutlich unwahrscheinlich ist, dass die Individualeffekte nicht mit den erklärenden Variablen korreliert sind.

Ein Test auf Gleichheit der random und der Fixed-Effects-Schätzergebnisse beruht auf dem allgemeinen Testprinzip von Hausman (1978).²⁶ Die Teststatistik lautet wie folgt:

$$(14) \quad \chi^2 = (\beta_{RE} - \beta_{FE})' (\Omega_{FE} - \Omega_{RE})^{-1} (\beta_{RE} - \beta_{FE}),$$

wobei β_{RE} und β_{FE} (Ω_{RE} und Ω_{FE}) die geschätzten Parametervektoren (Varianz-Kovarianzmatrizen) des Random bzw. des Fixed-Effects-Modells bezeichnen. Diese Teststatistik ist χ^2 verteilt mit Anzahl der Freiheitsgrade gleich dem Rang der Matrix $(\Omega_{FE} - \Omega_{RE})$.²⁷ Die Nullhypothese lautet, dass das Random-Effects-Modell korrekt sind, die individuellen Effekte also nicht mit den erklärenden Variablen korreliert sind. Die Teststatistik beruht auf dem Vergleich des unter der Nullhypothese konsistenten und effizienten, unter der Gegenhypothese jedoch inkonsistenten Random-Effects-Schätzers, mit einem unter Null- und Gegenhypothese konsistenten Fixed-Effects-Schätzer. Liegt die Teststatistik unter dem kritischen Wert aus der χ^2 -Verteilung, so kann das Random-Effects-Modell nicht abgelehnt werden.

25 Ein einfacher t-Test ist in diesem Falle nicht adäquat, weil das Random-Effects-Modell nicht nistend ist für das gepoolte Logitmodell.

26 Eine ausführliche Darstellung der Familie der Hausman-Tests geben Gourieroux und Montfort (1995, Kap. 18.4).

27 In der praktischen Anwendung taucht häufig das Problem auf, dass zeitinvariante Regressoren (z. B. Rechtsform oder Branchenzugehörigkeit) beim Fixed-Effects-Modell nicht berücksichtigt werden können. In diesem Falle müssen der Koeffizientenvektor und die Varianz-Kovarianzmatrix des Random-Effects-Modells entsprechend verkleinert werden.

Tabelle 6: Panel-Logit-Schätzergebnisse des Random-Effects-Modells

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	-4,4770	2,9888	0,134
ln(Umsatz) ²	0,1924	0,1302	0,139
Eigenkapitalquote	-7,2281	1,3809	0,000
Cash flow	-1,1850	0,5195	0,023
Anlagendeckungsgrad	-0,2709	0,2109	0,199
Beschr. Haftung	-0,0369	0,5483	0,946
1993	1,0812	0,4834	0,025
1994	2,0033	0,4704	0,000
1995	1,8346	0,4658	0,000
1996	2,2158	0,4749	0,000
1997	3,1193	0,5759	0,000
1998	3,4497	0,7330	0,000
Verarb. Gew.	-0,1441	0,5422	0,790
Baugewerbe	-1,3022	0,8159	0,110
Handel	-0,3510	0,6545	0,592
Sonstige	-0,0846	0,5391	0,875
Konstante	24,3907	16,9963	0,151
Corr.	2,2424	0,2996	0,000
Wald-Tests auf gemeinsame Signifikanz			
Variablen	chi²	d.o.f	p-Wert
Umsatz	2,31	2	0,31
Jahresdummies	43,98	6	0,00
Branchendummies	2,78	4	0,60
ges. Schätzung	84,82	16	0,00

Die Ergebnisse unterscheiden sich qualitativ nur sehr wenig von denen des gepoolten Probit- und geordneten Probit-Modells. Auch die Koeffizientenvektoren der verschiedenen Modelle sind sich, sofern die entsprechende Umrechnung der Probit- in Logitkoeffizienten vorgenommen wird, recht ähnlich. Dies überrascht deshalb, weil Vergleiche unterschiedlicher Logit- und Probitschätzungen streng genommen unzulässig sind und die Gleichheit der Koeffizientenvektoren von gepoolten und Random-Effects-Schätzergebnissen nur asymptotisch gilt und wir es hier mit einem relativ kleinen Datensatz zu tun haben.

Auffällig ist jedoch, dass die Präzision der Koeffizientenschätzung im random effects Fall erheblich zugenommen hat. Durch die Berücksichtigung der random effects können also beträchtliche Effizienzgewinne erzielt werden. Tatsächlich kann der Test auf random effects (Test auf $\sigma_u^2=0$) auf den konventionellen Signifikanzniveaus nicht abgelehnt werden. Dies bedeutet jedoch keineswegs, dass der Random-Effects-Schätzer deshalb dem Fixed-Effects-Schätzer vorzuziehen ist.

Mit einem Hausman-Test kann analysiert werden, ob ein „Poolen“ des Panel Datensatzes zulässig ist, also ob die Panelstruktur irrelevant ist oder ob ein einfaches binäres Logitmodell geschätzt werden sollte. Die korrespondierende Hausman-Teststatistik lautet:

$$(15) \quad \chi^2 = (\beta_{FE} - \beta)'(\Omega_{FE} - \Omega)^{-1}(\beta_{FE} - \beta),$$

wobei β und Ω die den Koeffizientenvektor und die Varianz-Kovarianzmatrix des gewöhnlichen Logitmodells bezeichnen. Die Nullhypothese lautet, dass keine unbeobachtbare Heterogenität vorliegt, ein Poolen des Datensatzes also zulässig ist. Die Teststatistik ist χ^2 -verteilt mit der Anzahl an Freiheitsgraden gleich Anzahl der erklärenden Variablen. In diesem Fall lautet der Testwert 15,83, das marginale Signifikanzniveau beträgt 0,199. Damit kann die Nullhypothese nicht verworfen werden. Ein Poolen des Datensatzes ist in diesem Fall also zulässig.

Tests auf Fehlspezifikation im binären Panelprobitmodell werden von Lechner (1995) vorgeschlagen. Auf eine Darstellung dieser Tests wird an dieser Stelle verzichtet, da es den Rahmen der Arbeit sprengen würde.

Sowohl für geordnete Wahrscheinlichkeitsmodelle als auch für binäre Probitmodelle existiert bislang nur ein Random-Effects-Schätzer, der in Tutz und Henevogl (1996) sowie Hamerle und Ronning (1995) beschrieben wird. Arulampalam (1999) nimmt zur Identifikation des geschätzten Koeffizientenvektors in Panel-Probitmodellen Stellung. Eine aktuelle Anwendung des geordneten random effects Probitmodells bieten Kaiser und Pfeiffer (2000).

Laisney und Lechner (1996) verwenden ein panel Probitmodell im Kontext einer GMM-Schätzung.²⁸ Ein wesentlicher Beitrag jenes Artikels ist es, dass gezeigt wird, dass die Verwendung von Makrovariablen als zusätzliche Momentenbedingungen zu deutlichen Effizienzgewinnen führt.

Als Gütemaße können sowohl für binäre als auch für geordnete Paneldatenmodelle die bereits in Abschnitt 3 besprochenen, auf Vergleichen der Werte der Log-Likelihoodfunktion beruhenden Pseudo-R²-Gütemaße verwendet werden.

3.5.2 *Multivariate Probitmodelle*

Beim multivariaten Probitmodell wird davon ausgegangen, dass die Fehlerterme von zwei (bivariate) oder mehr als zwei (multivariate) Probitgleichungen miteinander korreliert sind.²⁹ Eine getrennte Schätzung würde dann zu konsistenten, aber ineffizienten Schätzern führen.³⁰ Multivariate Probitmodelle tragen

28 GMM steht für Generalized Methods of Moments. Einen Überblick über GMM-Modelle bietet Mátyás (1999).

29 Bei multivariaten Logitmodellen ist die Kovarianzstruktur der einzelnen Gleichungen stark restringiert, so dass sie in der Praxis kaum angewendet werden (siehe Börsch-Supan, 1987).

30 Das Analogon multivariater Probitmodelle sind Modelle scheinbar unabhängiger Regressionen (seemingly unrelated regression, SUR) des linearen Regressionsmodells (siehe Greene, 1997, Kap. 15.4).

der Korrelation der Fehlerterme Rechnung und liefern somit Koeffizientenschätzungen mit höherer Präzision, also mit geringerer Standardabweichung, als separate Modelle.

Eine nahe liegende Erweiterung multivariater Probitmodelle sind geordnete multivariate Probitmodelle, wie sie von Kaiser (1999a) zur Analyse des Effektes neuer Technologien auf die Nachfrage nach verschiedenen qualifizierten Arbeitskräften angewendet werden.

Eine gerade für die Kreditpraxis relevante Erweiterung der multivariaten Probitmodelle liegt in den multivariaten Probitmodelle mit Selektionsmechanismus. So kann ein Kreditausfall z. B. nur dann beobachtet werden, wenn überhaupt ein Kredit vergeben wurde. Würde man also die Wahrscheinlichkeit eines Kreditausfalls lediglich mit den Informationen über diejenigen Kredite, die auch genehmigt worden sind, berechnen, so kann keine generelle Aussage über die Ausfallwahrscheinlichkeit eines Kreditnehmers gemacht werden. Es liegt ein Selektionsproblem vor – nur diejenigen Kreditnehmer, deren Anträge genehmigt wurden, sind in den Datensatz hineinselektiert worden.³¹ In diesen Fällen bietet sich das binäre Probitmodell mit Selektionskorrektur als adäquate Lösung des Problems an. Auch im geordneten Probitmodell kann für eine Selektionsverzerrung korrigiert werden. Kaiser (1999b) bietet eine Anwendung dieses Schätzansatzes.

Stichprobenselektionsverzerrungen können auch mit Hilfe des „choice-based“-Schätzers von Manski und Lerman (1977) in den Griff bekommen werden. Die Idee dieses Schätzers ist es, die einzelnen Beobachtungen innerhalb der Maximum-Likelihood-Schätzung gemäß ihrer Bedeutung in der entsprechenden Grundgesamtheit zu berücksichtigen. Boyes et al. (1989) wenden den Ansatz von Manski und Lerman (1977) auf das Kredit scoringproblem bei Kreditkartenbesitzern an. Da der Kreditanalyst nur dann Rückzahlungsschwierigkeiten beobachten kann, wenn überhaupt eine Kreditkarte vergeben wurde, beziehen sich Analysen von Rückzahlungsschwierigkeiten immer nur auf diejenigen, die eine Kreditkarte besitzen. Um aber Handlungsanweisungen dahingehend abgeben zu können, welche Personenkreise eine Kreditkarte erhalten sollten, muss diese Selektionsverzerrung korrigiert werden. Dazu muss die Grundgesamtheit aller potenziellen Kreditkartenbesitzer festgestellt werden. Diese würde dann in Personenkreise verschiedener Charakteristika (Alter, Beruf, Geschlecht etc.) eingeteilt. Die in den Kreditakten vorhanden Personen werden dann entsprechend ihrer – anhand der Charakteristika gemessenen – relativen Häufigkeit in der Gesamtpopulation auch bei der Schätzung gewichtet.

Auch der von uns hier verwendete Datensatz ist in Hinblick auf die Kreditvergabe verzerrt. Um das absolute, also nicht auf die Kreditvergabe bezogene, Kreditrisiko zu messen, müssten die Schätzungen entweder durch eine entsprechende

31 Im linearen Regressionsmodell können solche Selektionsprobleme durch die Einführung eines Korrekturterms behoben werden (siehe Verbeek, 2000, Kap. 7.5).

Selektionskorrektur oder einen choice-based sampling Ansatz angewendet werden.

3.5.3 *Simultane Probitmodelle*

Zwei (oder mehrere) zu analysierende Schätzgleichungen weisen häufig nicht nur eine gemeinsame Korrelationsstruktur der Fehlerterme auf, sondern sind auch direkt voneinander abhängig; ein Endogenitätsproblem tritt auf: Eine Variable X bestimmt die Variable Y und umgekehrt. In solchen Fällen werden die Modellparameter inkonsistent geschätzt. Das gilt sowohl für Probitmodelle als auch für das lineare Regressionsmodell.

So könnte man z. B. argumentieren, die Wahl der Rechtsform eines Unternehmens sei endogen für die Kreditausfallwahrscheinlichkeit dieses Unternehmens. Beispielsweise wird ein Unternehmer, der eine sehr riskante Geschäftsidee verfolgt, wohl kaum die Rechtsform einer Einzelfirma wählen, sondern eher eine GmbH gründen. Rechtsformwahl und Kreditausfallrisiko könnten also voneinander abhängig sein. Ein simultanes Probitmodell wäre dann das geeignete Analyseinstrument. In einer aktuellen Arbeit zum Zusammenhang zwischen Export und Innovation von Ebling und Janz (1999) wird ein solches simultanes Probitmodell angewendet. Eine allgemeine Darstellung finden sich in Maddala (1983), weitere Anwendungen und Verallgemeinerungen geben Pohlmeier und Entorf (1990) sowie Pohlmeier (1992).

4 *Übergangsratenmodelle*

Die bisher vorgestellten mikroökonomischen Modelle zur Analyse von Kreditausfallrisiken unterscheiden sich von den im folgenden beschriebenen Übergangsratenmodellen in zweierlei Hinsicht. Zum einen modellieren Übergangsratenmodelle die Wahrscheinlichkeit einer Zustandsveränderung direkt, während die in Kapitel 3 dargestellten Methoden letztlich auf ein latentes stetiges Modell zurückgreifen. Zum anderen berücksichtigen Logit- und Probitmodelle zusätzliche erklärende Variablen, während dies für Übergangsratenmodelle nicht der Fall ist. Sie spezifizieren die empirische Verteilung von Übergängen und Übergangswahrscheinlichkeiten.

Übergangsratenmodelle spezifizieren die Wahrscheinlichkeit einer Zustandsveränderung in Abhängigkeit vom Zustand in der Vorperiode (Markovketten) bzw. von der Verweildauer im gegenwärtigen Zustand (Kaplan-Meier-Schätzer). Mit Übergangsratenmodellen werden keine Koeffizienten (Parameter) geschätzt, sie werden daher auch als nichtparametrische Verfahren bezeichnet. Die in Kapitel 2 vorgestellten Ansätze werden hingegen als parametrisch bezeichnet.

4.1 Markovketten

Markovketten berechnen die Wahrscheinlichkeit des Überganges einer Variablen, im Anwendungsbeispiel eine Kredites, in einen oder mehrere andere Zustände.³² Die Übergangswahrscheinlichkeiten werden als konditional auf den Ursprungszustand gemessen:

$$(16) \quad P(K_{it} = l | K_{it-1} = m) = p_{im},$$

wobei m denselben oder einen anderen Zustand als l bezeichnet. Die Übergangswahrscheinlichkeit p_{im} bezeichnet also die Wahrscheinlichkeit, mit der Zustand l dem Zustand m folgt. Es wird folglich angenommen, dass der Zustand zum Zeitpunkt t lediglich vom Vorzustand zum Zeitpunkt $t-1$ abhängig ist. Dies ist die sogenannte „Markov-Eigenschaft“.³³ Die Übergangswahrscheinlichkeit wird berechnet als Quotient aus einerseits der Anzahl der Kredite, die sich in Zeitpunkt $t-1$ in Zustand m befunden haben *und* sich in Zeitpunkt t in Zustand l befinden sowie andererseits der Anzahl der Kredite, die sich in Zeitpunkt $t-1$ in Zustand m befunden haben. Eine häufige und übersichtliche Darstellung solcher Übergangswahrscheinlichkeiten stellen Übergangsmatrizen (transition matrices) dar. Für den Fall der drei Kreditzustände des geordneten Probitmodells aus Abschnitt 2 ergibt sich die Übergangsmatrix in Abbildung 8.

Tabelle 8: Übergangsmatrix

		Endzustand			
		kein Ausfall	teilweiser Ausfall	vollständiger Ausfall	
Start- zustand	kein Ausfall	453 84,83%	49 9,18%	32 5,99%	534
	teilweiser Ausfall	33 36,67%	38 42,22%	19 21,11%	90
	vollständiger Ausfall	13 16,46%	6 7,59%	60 75,95%	79
		499 70,98	93 13,23	111 15,79	703

Die Tabelle wird waagrecht (von links nach rechts) gelesen. Auf 84 Prozent der Beobachtungen, in denen keine Probleme sichtbar sind, folgt wieder eine Beobachtung ohne sichtbare Probleme, das heißt, die Unternehmen verlassen ihren

32 Eine kompakte und ausführliche Darstellung von Markovketten gibt Hamilton (1994, Kap. 22.2).

33 Es handelt sich hierbei um Markovketten erster Ordnung. Natürlich können auch Markovketten höherer Ordnung spezifiziert werden, allerdings erscheint es im Rahmen der Kreditrisikomessung als sinnvoll, von Markovketten erster Ordnung auszugehen.

Zustand nicht. In gut 9 Prozent der Fälle kommt es anschließend zu einem Teilausfall, in 6 Prozent der Fälle kommt es direkt zu einem Vollaussfall. Die geringen relativen Häufigkeiten (9 und 6 Prozent im Vergleich zu 85 Prozent) lassen sich durch die Schichtung des Datensatzes erklären. In etwa der Hälfte der hier beobachteten Unternehmen kommt es niemals zu Problemen. In der anderen Hälfte der Unternehmen tauchen irgendwann im beobachteten Zeitraum Probleme auf, ein Teil der Beobachtungen zeigt sich aber auch hier ohne sichtbare Zahlungsschwierigkeiten.

Auf den ersten Blick erstaunlich erscheinen die relativen Häufigkeiten in der zweiten Zeile. Sie zeigen an, dass es nach einem Teilausfall relativ häufiger zu einer Erholung kommt als zu einem Vollaussfall. Bei genauerer Analyse zeigt sich, dass es sich dabei oftmals um vorübergehende Erholungen beispielsweise nach der Verwertung von Sicherheiten oder der Stundung von Zins- und Tilgungszahlungen handelt. Solche vorübergehenden Erholungen können allein bei Betrachtung von Tabelle 8 nicht erkannt werden.

Bei der Analyse der untersten Zeile sei noch einmal daran erinnert, dass bei der Einteilung der Beobachtungen in die Klassen kein Ausfall, Teilausfall und Vollaussfall in die Klasse der Vollaussfälle auch Sanierungen gezählt wurden (siehe dazu Kapitel 2.2). In Fällen von erfolgreichen Sanierungen kann es im Anschluss an eine Beobachtung der Klasse Vollaussfall zu einer Verbesserung in eine bessere Klasse kommen, wie Tabelle 8 zeigt.

Bei all diesen Interpretationen muss bedacht werden, dass die Zahl der Beobachtung in den Kategorien unterschiedlich groß ist. Über den gesamten Beobachtungszeitraum hinweg weisen 703 Beobachtungen keine Probleme auf, nur 115 Beobachtungen zeigen Zahlungsschwierigkeiten und lediglich 126 erhebliche Probleme auf.

Eine weitere Schwierigkeit bei der Interpretation solcher einfacher Übergangsmodelle liegt darin, dass Heterogenitäten, also Unterschiede zwischen den Kreditnehmern nicht berücksichtigt werden. Daher bilden Ansätze eine Alternative, bei denen die Übergangswahrscheinlichkeit zwischen zwei oder mehreren Zuständen nicht nur vom jeweils vergangenen Zustand abhängt, sondern auch von zusätzlichen erklärenden Variablen. Einen solchen Ansatz beschreibt Gourieroux (1989). Im Kern handelt es sich dabei um ein multinomiales Logitmodell, also eine Erweiterung des in Kapitel 2 vorgestellten binären Logitmodells für mehrere ungeordnete Zustände. Nguyen Van et al. (2000) wenden dieses Verfahren an, um den Erfolg von Unternehmen des Dienstleistungssektors zu analysieren. Die Autoren erweitern zudem das Grundmodell von Gourieroux (1989) indem sie unbeobachtbare Heterogenität zulassen und die Effekte der Unternehmensgröße auf die Firmenperformance nichtparametrisch schätzen.

In einer aktuellen Arbeit zur Analyse von Firmenratings verwenden Nickell et al. (2000) geordnete Probitmodelle, um die Veränderung (Verbesserung, keine Veränderung, Verschlechterung) von Ratingentscheidungen abzubilden.

4.2 Kaplan-Meier-Schätzer

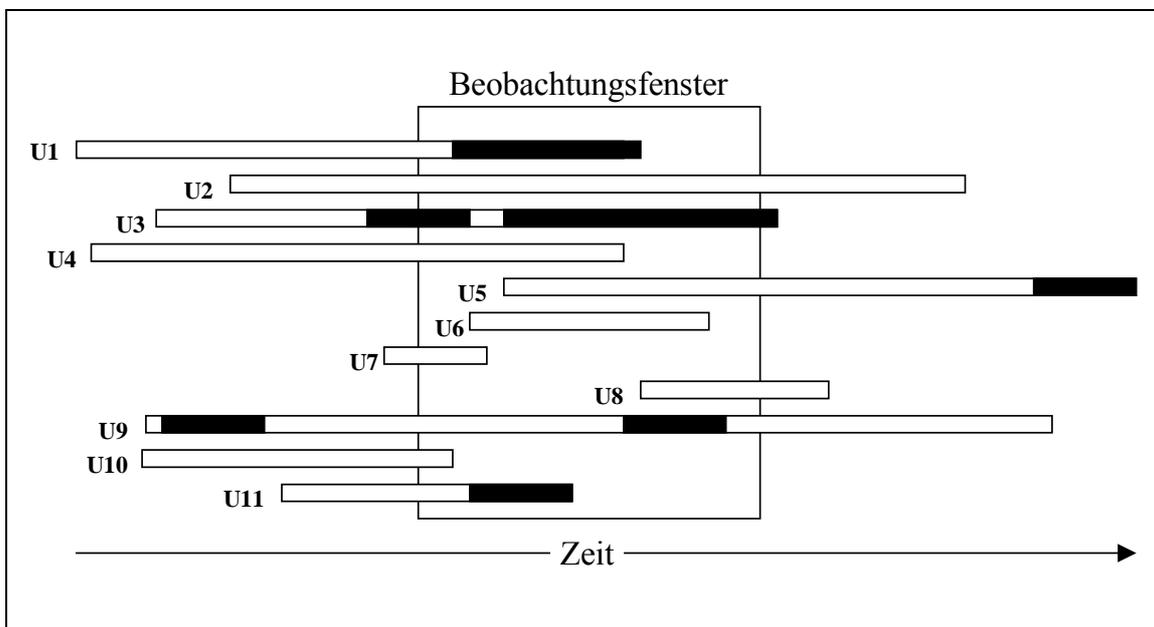
Während einfache Markovketten die Wahrscheinlichkeit eines Überganges von einem Kreditzustand in einen anderen in Anhängigkeit vom jeweils vorhergehenden Zustand berechnen, spezifiziert der Kaplan-Meier-Schätzer das Überleben in einem Zustand l , gegeben dass sich ein Kredit bereits bis zum Zeitpunkt t in diesem Zustand befunden hat. Im Gegensatz zum einfachen Markovkettenmodell lässt der Kaplan-Meier-Schätzer dabei nur zwei Zustände zu, z. B. Kreditausfall oder kein Kreditausfall. Der Kaplan-Meier-Schätzer berechnet also Überlebenswahrscheinlichkeiten. In Lehrbuch-Abhandlungen (z. B. Greene 1997, Kap. 20.5) wird der Kaplan-Meier-Schätzer lediglich für unzensierte Beobachtungen besprochen. Es wird also davon ausgegangen, dass eine Kreditverbindung über ihre gesamte Dauer beobachtet wird. In der Praxis ist dies selten der Fall, weil die Bereitstellung entsprechender Daten kostspielig und zeitaufwendig ist. Vielmehr gibt es in den Datensätzen, so auch in dem von uns verwendeten, viele Kreditengagements, die nicht bis zur vollständigen Rückzahlung beobachtet werden und/oder erst nach einigen Jahren des Bestehens der Kreditbeziehung in den Datensatz kommen. Im ersten Fall spricht man von einer „Rechtszensierung“, im zweiten von einer „Linkszensierung“. Bei einer Rechtszensierung kann die Kreditbeziehung nicht über einen bestimmten Zeitpunkt hinaus verfolgt werden. Diese Rechtszensierung muss in der Analyse berücksichtigt werden, weil ansonsten alle Kredite, die zum Ende der Analyseperiode noch nicht ausgefallen sind, als ausgefallen betrachtet würden. Das Ausfallrisiko würde in diesem Fall überschätzt. Bei dem in diesem Aufsatz verwendeten Datensatz ist dies in rund 70 Prozent der betrachteten Unternehmen der Fall. Ein umgekehrtes Problem ergibt sich bei Linkszensierungen. Linkszensierungen gibt es auch in dem von uns verwendeten Datensatz. Rund 12 Prozent der Unternehmen weisen bereits zum Anfangszeitpunkt Zahlungsschwierigkeiten auf. Für diese Unternehmen liegt keine Information vor, seit wann die Vertragsstörungen bestehen bzw. umgekehrt, wie lange es gedauert hat, bis erstmals ein Risiko aufgetreten ist.

Abbildung 1 zeigt Beispiele von Beobachtungsmustern, die unserem Datensatz entsprechen könnten. Der Kasten steht für den Beobachtungszeitraum – in unserem Fall sind es die Jahre 1992 – 1998. Die Querbalken U1 bis U11 sind Kreditengagements, die größtenteils vor Beginn des Beobachtungszeitraumes begonnen haben.

Mit Beginn des Kreditengagements sind die Unternehmen „im Risiko“. Im Datensatz finden sich nur wenige neue Kreditkunden, die erst während des Beobachtungszeitraumes Kreditkunden wurden. Beobachtungsmuster U5, U6 und U8 entsprechen solchen neuen Kreditkunden. Lediglich U6 ist ein Beispiel für ein Engagement, das vollständig beobachtet werden kann. Viele Unternehmen in unserem Datensatz beginnen und enden außerhalb des beobachteten Zeitraumes. U2, U3 und U9 sind Beispiele für solche Engagements. Schwarze Balken kennzeichnen aufgetretene Zahlungsschwierigkeiten. U9 beispielsweise hatte in der

Anfangsphase der Kreditbeziehung Zahlungsschwierigkeiten, von denen man keine Informationen im Datensatz finden kann. Einige Engagements fallen während des Beobachtungszeitraumes aus, so etwa U1 und U11. Andere Engagements enden vertragsgemäß mit der Rückzahlung des Kredites bzw. der Kredite. U4, U6, U7 und U10 zeigen solche Beobachtungsmuster. Von U2, U3, U5, U8 und U9 konnte man bei Abschluss der Beobachtungszeit noch nicht sagen, ob die Kreditverträge erfüllt werden oder nicht. U1 bis U4, U7 und U9 bis U11 sind linkszensiert, da sie bereits vor Beginn des Beobachtungszeitraumes existierten. U2, U5 und U9 werden als rechtszensiert bezeichnet, da sie nach Ende des beobachtungszeitraumes noch weiter existieren.

Abbildung 1: Beobachtungsmuster von Kreditengagements im Datensatz



U1 bis U11 sind Beispiele für Beobachtungsmuster von Kreditengagements mit Unternehmen. Das Beobachtungsfenster kennzeichnet den Zeitraum, in dem Daten aus den Kreditengagements erhoben wurden. Die Balken stehen für das Bestehen des Kreditengagements, schwarze Abschnitte kennzeichnen Zahlungsschwierigkeiten.

Die im folgenden vorgestellten Modelle zur Verweildaueranalyse werden zur Vereinfachung dennoch dargestellt, als wenn keine Zensierungsprobleme vorlägen.³⁴ Die von uns vorgestellten empirischen Ergebnisse tragen den Zensierungsproblemen jedoch Rechnung. In den einzelnen Unterabschnitten geben wir Hinweise auf Referenzen, die auf die Zensierungsproblematik eingehen.

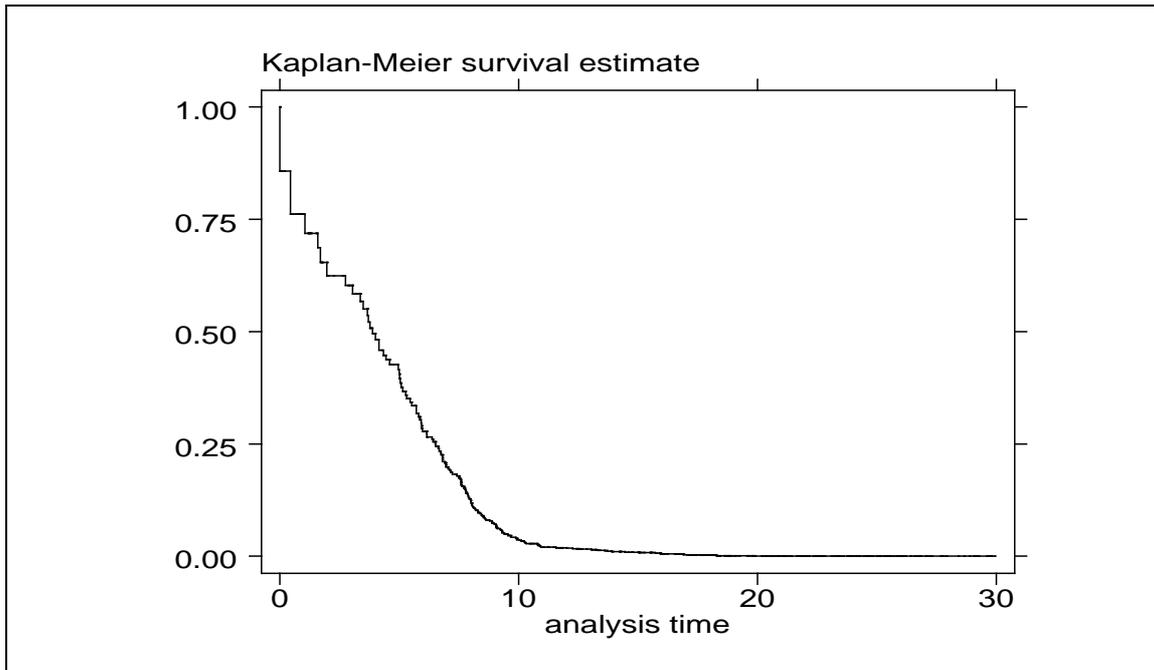
Der Kaplan-Meier Schätzer der Überlebenswahrscheinlichkeit in einem Zustand – im konkreten Anwendungsfall die Wahrscheinlichkeit eines Nicht-Ausfalles der Kreditrückzahlung – berechnet sich als

³⁴ Dies mag in der Praxis auch deshalb gerechtfertigt sein, weil Standardsoftwareprogramme wie z.B. STATA Zensierungsprobleme automatisch korrigieren, sofern ihnen die entsprechenden Zensierungsinformationen gegeben werden.

$$(18) \quad \hat{S}(T_k) = \prod_{i=1}^k \frac{n_i - h_i}{n_i},$$

wobei n_i die Anzahl der Kredite bezeichnet, die mindestens bis zum Zeitpunkt T_k (bis zur Dauer T_k der Vertragsbeziehung) überlebt haben und h_i die Anzahl der Kredite bezeichnet, die zum Zeitpunkt T_k ausgefallen sind.

Abbildung 2: Nichtausfallwahrscheinlichkeit



Dargestellt wird die Nichtausfallwahrscheinlichkeit von Krediten. Je länger das Engagement existiert, desto geringer ist die Nichtausfallwahrscheinlichkeit, die auf der Skala von 0.00 (= 0 Prozent) bis 1.00 (= 100 Prozent) angegeben wird. Die Analysezeit ist in Jahren dargestellt.

Abbildung 2 bildet die empirische Überlebenswahrscheinlichkeit, d. h. die empirische Nichtausfallwahrscheinlichkeit der Kredite in unserem Datensatz ab.³⁵ Diese Abbildung entspricht einer graphischen Übersetzung einer aus der Versicherungsmathematik bekannten Sterbetabelle.

Zwar mag die Nichtausfallwahrscheinlichkeit eines Kredites in vielen Fällen von Interesse sein, noch relevanter ist jedoch Information darüber, ob ein Kredit sehr bald ausfällt, gegeben dass er bis zum Zeitpunkt T_k nicht ausgefallen ist. Die Wahrscheinlichkeit, dass eine Verweildauer in der nächsten kurzen Periode endet wird als Hazardrate $\lambda(T_k)$ bezeichnet. Für den Kaplan-Meier-Schätzer ist sie definiert als

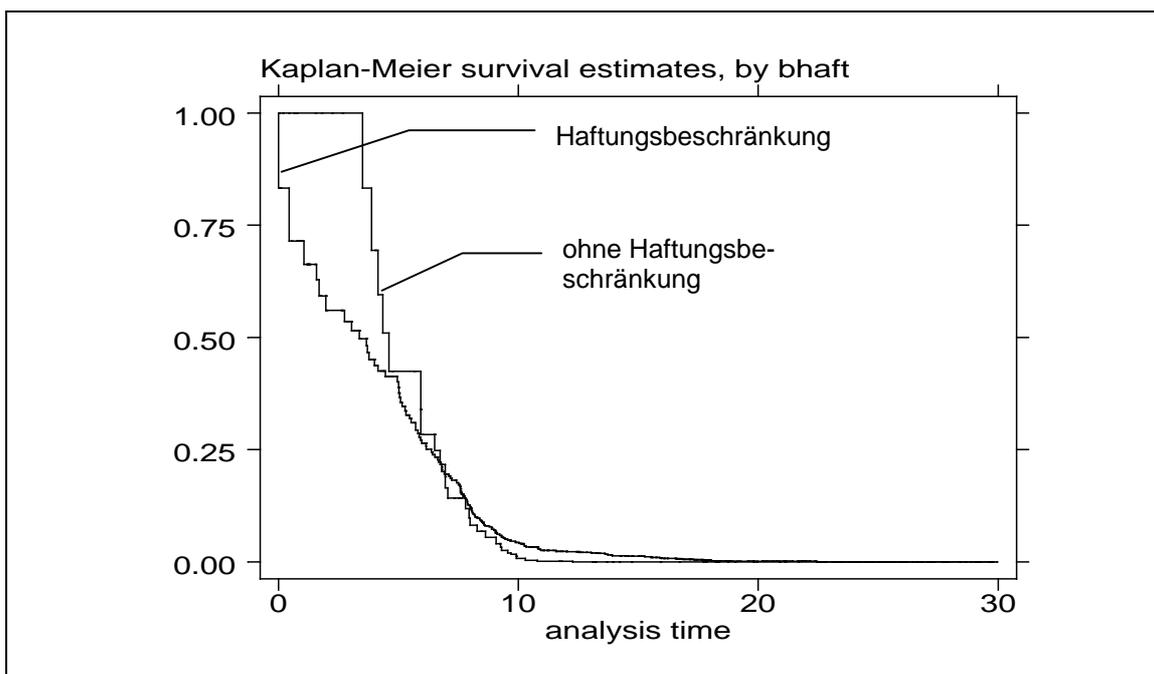
$$(19) \quad \lambda(T_k) = \frac{h_k}{n_k}$$

³⁵ Auch eine Schätzung der Standardfehler der Überlebenswahrscheinlichkeiten ist möglich, siehe Kalbfleisch und Prentice (1980, Kap. 1.3)

und entspricht dem Verhältnis von gerade ausgefallenen zu noch nicht ausgefallen – sich „im Risiko“ befindlichen – Krediten. Der Kaplan-Meier Schätzer entspricht der Überlebenswahrscheinlichkeit, nicht aber der Hazardrate, hat also immer einen fallenden Verlauf.

Zur Überprüfung von Unterschieden verschiedener Risikogruppen kann ein log-rank oder Gehan-Whitney Test herangezogen werden (Kalbfleisch und Prentice 1980, S. 16f. oder Blossfeld et al. 1986, S. 46f.). Geht man z. B. davon aus, dass Kredite einer Gesellschaft mit beschränkter Haftung weniger lange ohne Risiko bleiben als Kredite unbeschränkt haftender Gesellschafter, so kann diese These mit den beiden Tests überprüft werden. Für den hier verwendeten Datensatz ergibt sich, dass Unternehmen mit Haftungsbeschränkung früher ausfallen als Unternehmen ohne Haftungsbeschränkung. Die Teststatistik hat den Wert 4,14, das marginale Signifikanzniveau ist 0,0418, die beiden Überlebensfunktionen unterscheiden sich also signifikant voneinander.

Abbildung 3: Nichtausfallwahrscheinlichkeit in Abhängigkeit von der Rechtsform



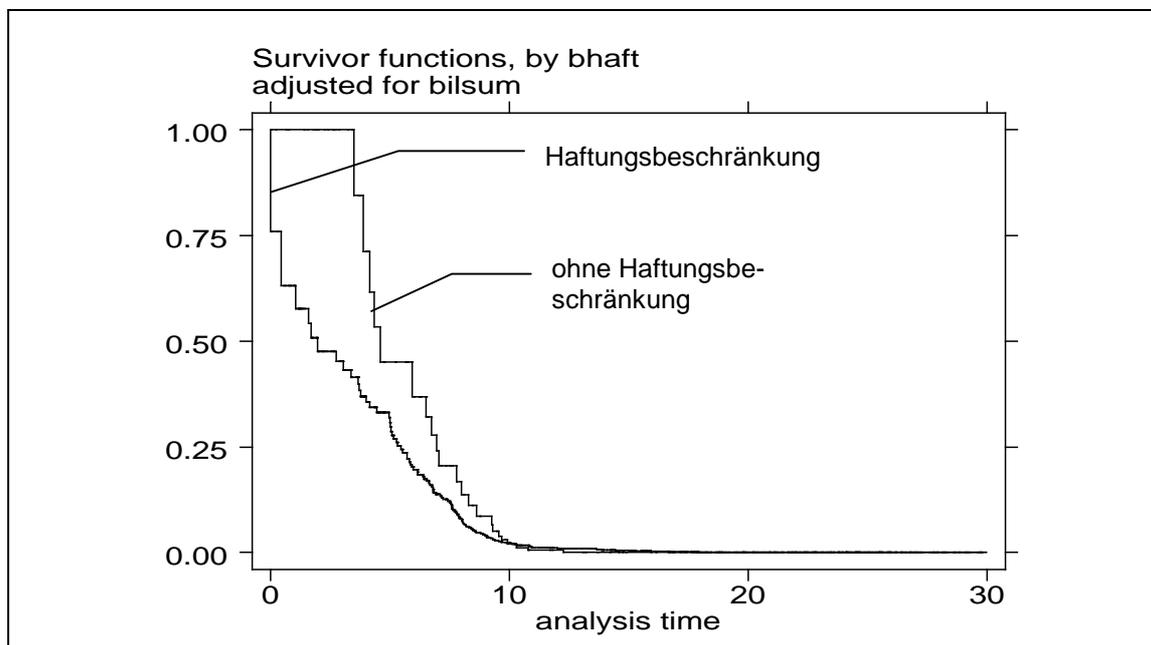
Die Nichtausfallwahrscheinlichkeit wird in Abhängigkeit von der Haftungsbeschränkung (by bhaft) der betrachteten Unternehmen dargestellt. Je länger das Engagement existiert, desto geringer ist die Nichtausfallwahrscheinlichkeit, die auf der Skala von 0,00 (= 0 Prozent) bis 1,00 (= 100 Prozent) angegeben wird. Die Analysezeit ist in Jahren dargestellt.

Abbildung 3 zeigt die unterschiedlichen Nichtausfallwahrscheinlichkeiten für Unternehmen mit und ohne Haftungsbeschränkung. Es wird deutlich, dass Unternehmen mit Haftungsbeschränkung tendenziell früher ausfallen als Unternehmen ohne Haftungsbeschränkung. Nach Stiglitz und Weiss (1981) weist eine

erhöhte Haftungsbeschränkung indirekt auf den Unwillen eines Unternehmens hin, das Risiko seiner Geschäftstätigkeit alleine zu tragen, womit sich im Schnitt ein risikofreudigeres Verhalten erwarten lässt. Ein risikofreudigeres Verhalten ist allerdings auch mit einer höheren Ausfallquote verbunden.

Die beiden Kurven in Abbildung 3 schneiden sich jedoch, was vermuten lässt, dass hier ein weiterer Effekt den Zusammenhang zwischen Haftung und Nichtausfallwahrscheinlichkeit beeinflusst, ohne explizit betrachtet zu werden. So sind Unternehmen mit einer haftungsbeschränkten Rechtsform wie der GmbH oder der AG im Schnitt größer als Unternehmen, deren Unternehmer mit ihrem persönlichen Vermögen haften. Größere Unternehmen weisen etwa durch ihre höhere Marktmacht und größere Diversifikationsmöglichkeiten eine geringere Ausfallwahrscheinlichkeit auf. Um solche störenden Einflüsse bei der Betrachtung von Nichtausfallwahrscheinlichkeiten in unterschiedlichen Unternehmensgruppen verringern zu können, kann man die Kaplan-Meier-Schätzung korrigieren. In diesem Fall wurde mit Hilfe der Bilanzsumme als Größenindikator ein möglicher Größeneinfluss korrigiert. Abbildung 4 zeigt als Ergebnis eine klarere Unterscheidung zwischen Unternehmen mit und ohne Haftungsbeschränkung.

Abbildung 4: Nichtausfallwahrscheinlichkeit in Abhängigkeit von der Rechtsform



Die Nichtausfallwahrscheinlichkeit wird – unter Berücksichtigung der Unternehmensgröße (adjusted for bilsum) – in Abhängigkeit von der Haftungsbeschränkung (by bhaft) der betrachteten Unternehmen dargestellt. Je länger das Engagement existiert, desto geringer ist die Nichtausfallwahrscheinlichkeit, die auf der Skala von 0,00 (= 0 Prozent) bis 1,00 (= 100 Prozent) angegeben wird. Die Analysezeit ist in Jahren dargestellt.

Das beschriebene Beispiel zeigt, dass ein grundlegender Nachteil des Kaplan-Meier-Schätzers darin liegt, dass beobachtbare Unterschiede zwischen den Kreditnehmern mehr oder weniger nur univariat berücksichtigt werden können. Insofern ist der nahe liegende nächste Schritt, Verweildauermodelle zu verwenden, die erklärende Variablen berücksichtigen können. Solche Modelle werden in den beiden folgenden Abschnitten dargestellt.

5 Verweildauermodelle

5.1 Modelle proportionaler Hazardraten: parametrische Modelle

Eine grundlegende Frage bei der Spezifikation von parametrischen Verweildauermodellen ist es, welchen Verlauf die Hazardrate hat. Ist es plausibel, dass die Hazardrate, also im konkreten Anwendungsfall die Gefahr eines Kreditausfalls, im Zeitablauf abnimmt? Dass sie zunimmt? Oder ist sie völlig unabhängig von der Verweildauer im Nichtausfallzustand? Die Ergebnisse parametrischer Verweildauermodelle hängen in hohem Maße von der korrekten Spezifikation des Hazardratenverlaufes ab. Ist der Verlauf der Hazardrate z. B. real ansteigend, der Analyst lässt aber lediglich konstante Hazardraten zu, dann kommt es zu einer Fehlspezifikation des Modells. Insofern bietet es sich an, vorab mögliche Verläufe der Hazardrate zu überprüfen, in gewisser Weise also die Daten selbst bestimmen zu lassen, welche funktionale Spezifikation der Hazardrate zu wählen ist.

Einen groben Anhaltspunkt dafür, welche funktionale Form für die Spezifikation der Hazardrate sinnvoll ist, bietet ein Plot der mit Hilfe des Kaplan-Meier Schätzers berechneten integrierten Hazardrate $\hat{\Lambda}$, die selbst keine inhaltliche Interpretation besitzt und sich als $\Lambda(\hat{T}_k) = \sum_{i \leq k} \hat{\lambda}(T_i)$ berechnet, gegen die einzelnen

Verweildauern. Ist der Verlauf der integrierten Hazardrate über die Verweildauern konvex (konkav), so impliziert dies eine zunehmende (abnehmende) Hazardrate. Man spricht auch von positiver (negativer) *duration dependence*. Ist der Verlauf der integrierten Hazardrate linear, so ist die Hazardrate Verweildauerunabhängig oder *memoryless*.³⁶

Ein solcher Plot der integrierten Hazardrate für die Kreditdaten unserer Anwendung ergibt einen leicht konvexen Verlauf über die Zeit und impliziert daher eine zunehmende Hazardrate. Insofern ist es nahe liegend, die Hazardrate mit einer Verteilung abzubilden, die zunehmende Hazardraten zulässt. Eine solche Verteilungsfunktion ist die Weibull-Verteilung, für die die Hazardrate folgende Form annimmt: $\lambda(t) = \gamma \alpha t^{\alpha-1}$. Die korrespondierende Hazardrate ist in Abhän-

³⁶ Natürlich könnte man im Prinzip auch direkt die geschätzten Hazardraten gegen die Verweildauern plotten. Aufgrund der hohen Erratik dieser Plots ist es jedoch schwierig, aus ihnen Informationen über den Verlauf der Hazardrate zu gewinnen.

gigkeit von der Größe von α steigend, konstant oder fallend. Für $\alpha > 1$ ist die Hazardrate steigend, für $\alpha < 1$ ist sie fallend und für $\alpha = 1$ ist sie konstant.³⁷

Modelle der proportionalen Hazardrate machen die Hazardrate von einem Vektor erklärender Variablen x_i abhängig, einem korrespondierenden Koeffizientenvektor β und einem „baseline hazard“, einer Grundhazardrate $\lambda_0(t_k)$, die über alle Individuen (über alle Kredite) als konstant angenommen wird. Die Hazardfunktion wird also ausgedrückt als:

$$(20) \quad \lambda(t_k, x_k, \beta, \lambda_0) = \phi(x_k, \beta) \lambda_0(t_k).$$

Der Ausdruck $\phi(x_k, \beta)$ repräsentiert dabei die Heterogenität der Individuen. Beim „baseline hazard“ $\phi(x_k, \beta)$ handelt es sich um eine Funktion, die im Prinzip für jedes Individuum geschätzt werden muss. Der Effekt der erklärenden Variable ist es, die Grundhazardrate mit einem Faktor $\phi(\cdot)$ zu multiplizieren.³⁸ Die Funktion $\phi(x_k, \beta)$ wird häufig als Exponentialfunktion, $\phi(x_k, \beta) = \exp(x_k \beta)$, definiert.

Um das Modell identifizieren zu können, muss – ähnlich wie bei den binären und den geordneten Modellen – eine Normierung auferlegt werden.³⁹ Für das Exponentialmodell wird die Normierung $\lambda(t_k) = 1$ verwendet, für das Weibullmodell gilt $\lambda(t_k) = t^{\alpha-1}$. Beide Normierungen ergeben sich aus einer Evaluation der Hazardrate an der Stelle $x_k = 0$.

Der Effekt der erklärenden Variablen ist es, die Grundhazardrate mit einem Faktor $\phi(\cdot)$ zu multiplizieren.⁴⁰ Die Funktion $\phi(x_k, \beta)$ wird häufig als Exponentialfunktion, $\phi(x_k, \beta) = \exp(x_k \beta)$, definiert. Daraus ergibt sich, dass der marginale Effekt der m ten erklärenden Variable auf die logarithmierte Hazardrate einfach dem Koeffizienten dieser erklärenden Variable entspricht:⁴¹

37 Für $\alpha = 1$ ergibt sich die Exponentialverteilung der Hazardrate.

38 Dieser Faktor $\phi(\cdot)$ selbst wird als zeitinvariant angenommen. In der Praxis ist diese Annahme, dass die erklärenden Variablen sich über die Verweildauern nicht verändern allerdings häufig verletzt, z.T. auch im hier vorliegenden Datensatz. Die Zeitabhängigkeit der Regressoren kompliziert das Schätzverfahren erheblich, und ein Eingehen auf die hier adäquaten Modelle würde den Rahmen dieses Aufsatzes sprengen. Daher kann an dieser Stelle lediglich auf das Lehrbuch von Lancaster (1990, Kap. 2.3) und auf Jenkins (1995) verwiesen werden.

39 Für beide Modelle bedeutet dies, dass direkte Koeffizientenvergleiche sowohl zwischen verschiedenen binären oder geordneten Probit/Logitmodellen als auch zwischen den Ergebnissen zweier oder mehrerer proportionaler Hazardratenmodelle unzulässig sind.

40 Dieser Faktor $\phi(\cdot)$ selbst wird als zeitinvariant angenommen. In der Praxis ist diese Annahme, dass die erklärenden Variablen sich über die Verweildauern nicht verändern allerdings häufig verletzt, z.T. auch im hier vorliegenden Datensatz. Die Zeitabhängigkeit der Regressoren kompliziert das Schätzverfahren erheblich, und ein Eingehen auf die hier adäquaten Modelle würden den Rahmen dieses Aufsatzes sprengen. Daher kann an dieser Stelle lediglich auf das Lehrbuch von Lancaster (1990, Kap. 2.3) und auf Jenkins (1995) verwiesen werden.

41 Genau wie im linearen Regressionsmodell repräsentieren also die Koeffizienten logarithmierter Variablen Elastizitäten. Koeffizienten, die zu Dummy-Variablen gehören, repräsentieren relative Veränderungen – relativ zur jeweiligen Basiskategorie – der Hazardrate.

$$(20) \quad \frac{\partial (\ln(\lambda(t_k, x_k, \beta, \lambda_0)))}{\partial x_{km}} = \beta_m.$$

Zur Schätzung des Effektes der erklärenden Variablen auf die Hazardrate ist es also nicht notwendig, die Grundhazardrate zu identifizieren.

Der Name „proportional hazard“ ergibt sich daraus, dass die Hazardraten zweier Individuen proportional zueinander sind, dieses Verhältnis ist also zeitunabhängig, weil sich die Grundhazardrate bei der Verhältnisbildung herauskürzt:

$$(21) \quad \frac{\lambda(t_k, x_k, \beta, \lambda_0)}{\lambda(t_j, x_j, \beta, \lambda_0)} = \exp((x_k - x_j)\beta).$$

D.h., dass bei der Einbeziehung von zeitinvarianten erklärenden Variablen wie z. B. Rechtsform sich das Verhältnis der Hazardraten von beschränkt und vollständig haftenden Kreditnehmern nicht verändern darf. Diese Annahme kann durch die Einführung von subpopulationsspezifischen Grundhazardraten, in diesem Falle bezogen auf die Rechtsform, gelockert werden.

Parametrische Verweildauermodelle wie das Weibull- oder das Exponentialmodell werden mit dem Maximum-Likelihood-Verfahren geschätzt. Die Herleitung der Likelihoodfunktion ist in Kiefer (1988, Kap. B) angegeben.

Tabelle 9 zeigt die Weibull-Schätzergebnisse für die Verweildauer im Nicht-Kreditausfallzustand. Als erklärende Variablen wurden alle bereits in den vorangegangenen Modellen verwendeten Variablen berücksichtigt. Nicht berücksichtigt wurden jedoch die Jahresdummyvariablen. Dies liegt daran, dass unser Datensatz fast ausschließlich Beobachtungen erhält, die allesamt bereits zum ersten Beobachtungszeitpunkt vorliegen. In der Terminologie von Lancaster (1990, Kapitel 8) liegen hier also „stock“ (Bestands-) Daten im Gegensatz zu „flow“ (Strom-) Daten vor. Bei einer Schätzung auf Grundlage von Bestandsdaten würden Zeitdummies daher direkt die Zeitabhängigkeit der Hazardrate abbilden.

Wie bereits bei den im dritten Abschnitt besprochenen Modellen, zeigen die Eigenkapitalquote und der dynamische Cash Flow einen signifikanten Einfluss auf das Kreditrisiko. Beide haben einen negativen Einfluss auf die Hazardrate und somit einen positiven Effekt auf die Überlebenswahrscheinlichkeit. Da Tabelle 9 die Schätzung der logarithmierten Hazardrate abbildet, beträgt der marginale Effekt des Cash flow (der Eigenkapitalquote) auf die Hazardrate -3,7 (-0,8) Prozent. Entsprechend lauten die Effekte auf die Überlebenswahrscheinlichkeit $\exp(-3,7)=0,02$ bzw. $\exp(-0,8)=0,4$. Die Überlebenswahrscheinlichkeit eines Kredites erhöht sich also um 0,98 Prozentpunkte bei einer einprozentigen Erhöhung des Cash flow. Bei einer einprozentigen Erhöhung der Eigenkapitalquote erhöht sich die Überlebenswahrscheinlichkeit um 0,6 Prozentpunkte.

Tabelle 9: Weibull-Schätzergebnis

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	0,7260	0,9439	0,442
ln(Umsatz) ²	-0,0379	0,0417	0,364
Eigenkapitalquote	-3,7136	0,4542	0,000
Cash flow	-0,7900	0,2408	0,001
Anlagendeckungsgrad	-0,0076	0,0229	0,738
Beschr. Haftung	-0,0253	0,1673	0,880
Verarb. Gew.	0,0103	0,1689	0,952
Baugewerbe	0,0332	0,2509	0,895
Handel	-0,1860	0,2099	0,375
Sonstige	0,1388	0,1663	0,404
Konstante	-4,4862	5,3253	0,400
	1,2058	0,0699	0,000
Wald-Tests auf gemeinsame Signifikanz			
Variablen	chi²	d.o.f	p-Wert
Umsatz	4,54	2	0,10
Branchendummies	2,63	4	0,62
ges. Schätzung	151,60	10	0,00

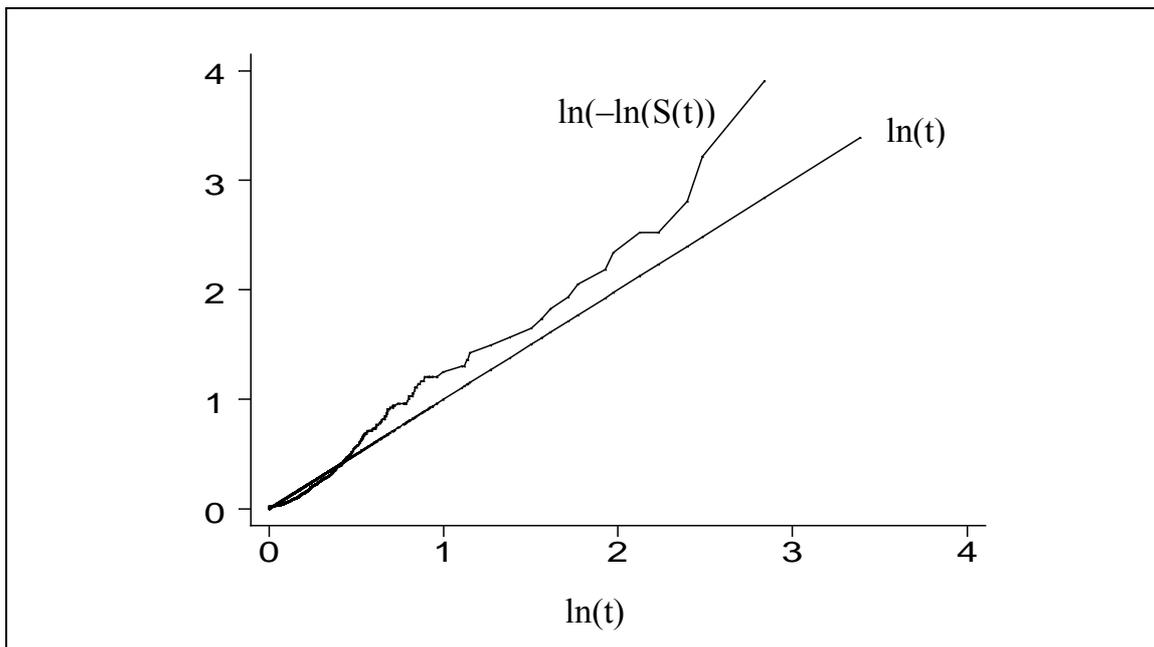
Der Schätzwert für den Parameter α , von dessen Größe der Verlauf der Hazardrate abhängt, ist 1,2148 mit einem Standardfehler von 0,0573 und ist damit hoch signifikant größer als 1. Evidenz für eine positive Verweildauerabhängigkeit liegt also vor: Je länger der Zeitraum ist, in der ein Kredit nicht ausgefallen ist, desto höher ist die Wahrscheinlichkeit, dass er in Kürze ausfallen wird.

Kalbfleisch und Prentice (1980, S. 24f) schlagen einen empirischen Test auf die Richtigkeit der Weibull-Verteilung vor. Die Überlebensfunktion der Weibull-Verteilung ist $S(t) = \exp(-\gamma t^\alpha)$ so dass gilt:

$$(22) \quad -\ln\left(\hat{S}(t)\right) = \gamma t^\alpha \quad \text{und} \quad \ln\left(-\ln\left(\hat{S}(t)\right)\right) = \ln(\gamma) + \alpha \ln(t),$$

wobei $\hat{S}(t)$ die Kaplan-Meier Schätzung der Überlebensfunktion bezeichnet. Ein Plot von $\ln\left(-\ln\left(\hat{S}(t)\right)\right)$ gegen $\ln(t)$ sollte daher eine gerade Linie mit der Steigung α ergeben. Ein solcher diagnostischer Plot ist in Abbildung 3 dargestellt.

Abbildung 5: Diagnostischer Plot



Eine lineare Regression von $\ln(t)$ auf $\ln\left(-\ln\left(\hat{S}(t)\right)\right)$ unterstützt die graphische Analyse. Der geschätzte Wert von α beträgt 1,2342 und ist hochsignifikant größer als 1.

In einem aktuellen Beitrag stellen van den Berg und van der Klaauw (1998) ein Bayesianisches Schätzverfahren für Verweildauermodelle vor, das den Einfluss von Makrodaten auf Dauer von Arbeitslosigkeit mitberücksichtigt.⁴² Die Berücksichtigung solcher Makrovariablen, z. B. der konjunkturellen Lage oder der Zahl der Konkurse, ist für die Kreditrisikomessung von potenziell hoher Bedeutung.

5.2 Modelle proportionaler Hazardraten: semiparametrische Modelle

Ein Nachteil der parametrischen Hazardratenmodelle ist es, dass die Grundhazardrate nicht vollständig identifiziert wird. Der Partial-Likelihood-Schätzer von Cox (1972, 1975) stellt daher eine Methode dar, den Koeffizientenvektor β zu schätzen, ohne Annahmen über die Grundhazardrate zu treffen.

Wenn die einzelnen Verweildauern – also die Zeitintervalle, in denen kein Kredit ausgefallen ist – ihrer Länge nach geordnet sind, dann ist die konditionale Wahrscheinlichkeit – konditional eben darauf, dass ein Kredit bis zu diesem Zeitpunkt noch nicht ausgefallen ist –, dass ein Kredit zum Zeitpunkt t_k ausfällt, wobei es sich hier um den ersten ausgefallenen Kredit handelt:

⁴² Eine Einführung in die bayesianische Ökonometrie bieten Gourieroux und Montfort (1995, Kap. 4).

$$(23) \quad \frac{\lambda(t_l, x_l, \beta)}{N} \cdot \prod_{k=1} \lambda(t_l, x_k, \beta)$$

Unter der Proportional-Hazard-Annahme $\lambda(t, x, \beta, \lambda_0) = \phi(x, \beta) \lambda_0(t)$ vereinfacht sich dieser Ausdruck zu

$$(24) \quad \frac{\phi(t_l, x_l, \beta)}{N} \cdot \prod_{k=1} \phi(t_l, x_k, \beta)$$

Unter Abwesenheit von Information über die Grundhazardrate spielt also nur die Ordnung der Verweildauern bei der Bestimmung der Hazardrate eine Rolle.⁴³

Das Modell von Cox wird vielfach als vertretbarer Kompromiss zwischen dem nichtparametrischen Kaplan-Meier Schätzer und möglicherweise vollkommen überparametrisierten parametrischen Modellen gesehen (z. B. Greene 1997, S. 999). Ein wesentlicher Nachteil des Modells von Cox ist jedoch, dass weder Überlebens- noch Hazardfunktionen berechnet werden können.

Tabelle 10: Partial-Likelihood-Schätzergebnis

Variable	Koeff.	Std.fehler	p-Wert
ln(Umsatz)	0,7358	0,9778	0,452
ln(Umsatz) ²	-0,0374	0,0431	0,385
Eigenkapitalquote	-3,6733	0,4976	0,000
Cash flow	-0,9637	0,2696	0,000
Anlagedeckungsgrad	0,0092	0,0232	0,692
Beschr. Haftung	-0,0867	0,1834	0,636
Verarb. Gew.	0,0224	0,1801	0,901
Baugewerbe	0,0826	0,2589	0,750
Handel	-0,2442	0,2244	0,277
Sonstige	0,0544	0,1735	0,754
Wald-Tests auf gemeinsame Signifikanz			
Variablen	chi²	d.o.f	p-Wert
Umsatz	3,32	2	0,19
Branchendummies	2,16	4	0,71
ges. Schätzung	127,23	10	0,00

Ein weiteres semiparametrisches Modell wurde von Han und Hausman (1990) vorgeschlagen. Dabei wird die Wahrscheinlichkeit, dass eine Verweildauer eine

⁴³ Als Konsequenz der Wichtigkeit der Ordnung der Verweildauern ergibt sich natürlich dann ein Problem, wenn viele Verweildauern gleicher Länge vorliegen. In der Kreditpraxis wird dieses Phänomen jedoch selten vorkommen und kann ohnehin unter bestimmten Bedingungen durch Gewichtung einzelner Verweildauern korrigiert werden (siehe Kalbfleisch und Prentice, 1980: 74f).

bestimmte Periode andauert, mit einem geordneten Probitmodell geschätzt. Eine Anwendung dieses Verfahrens ist nur dann empfehlenswert, wenn die Verweildauern diskret sind und viele Verweildauern gleicher Länge vorliegen.

Als Gütemaße können die bereits in Abschnitt 3 besprochenen Gütemaße verwendet werden.

Die Ergebnisse der semiparametrischen Schätzung unterscheiden sich nicht stark von denen der Weibull-Schätzung.

Die sowohl in den parametrischen als auch semiparametrischen Hazardraten-Modellen explizit modellierte Zeitdimension kann für das Kreditrisikomanagement sehr hilfreich sein. Die Informationen, die aus solchen Schätzmodellen gewonnen werden, könnten etwa für eine Laufzeitoptimierung von Kreditverträgen oder auch die Abstimmung von Hedging-Strategien genutzt werden.

5.3 *Weiterführende Literatur*

In diesem Abschnitt wurden Grundideen von Verweildauermodellen vereinfacht und in kurzer Form dargestellt. Eine hervorragende und umfangreichere Übersicht über Verweildauermodelle bietet ein Artikel von Kiefer (1988). Ein sehr empfehlenswertes Lehrbuch ist die Monographie von Lancaster (1990), der zahlreiche Erweiterungen – z. B. die Berücksichtigung von unbeobachtbarer Heterogenität der Individuen – bespricht und auch ausführlich auf Spezifikations-tests eingeht. Ein empfehlenswertes Einführungslehrbuch ist Blossfeld und Rohwer (1995).

6 *Weitere mikroökonomische Modelle*

Die in Kapitel 5 besprochenen Modelle können nur dann angewendet werden, wenn es lediglich zwei Risikozustände, z. B. Ausfall oder Nichtausfall, gibt. Eine nahe liegende Erweiterung sind daher Modelle, die es erlauben, mehrere Risiken zu modellieren. Solche *Modelle konkurrierender Risiken* („competing risk“) werden in Lancaster (1990, Kap. 5.5) ausführlich beschrieben. Eine Anwendung auf Marktaustritte von Firmen bieten Harhoff et al. (1998). Dolton und van der Klaauw (1999) beschäftigen sich mit dem Erwerbsverhalten von Lehrern und verwenden dabei eine Erweiterung des Ansatzes Han und Hausman (1990) für competing risk Modelle. Spezifikationstests für das „competing risk“ Modell werden von Pudney und Thomas (1995) vorgestellt und auf Verweildauern in Arbeitslosigkeit angewendet.

Während Verweildauermodelle die Zeit bis zum Eintritt eines bestimmten Ereignisses, hier das eines Kreditausfalls, analysieren, untersuchen *Zählmodellen*, wie viele dieser Ereignisse in einem bestimmten Zeitraum auftreten und welche erklärende Variablen das Auftreten solcher Ereignisse in welcher Weise beeinflussen. Zählmodellen basieren auf Annahmen über die Ereignis-Prozesse. In der Praxis werden häufig das Poisson-Modell und die hinsichtlich

ihrer funktionalen Form flexibleren Negbin-Modelle verwendet. Während erstere auf einem Poisson-Prozess basieren, beruhen letztere auf der Negativbinomialverteilung und nisten den Poisson-Prozess als Spezialfall.

Obwohl Zählmodellen für die Kreditvergabepraxis von geringer Bedeutung sind, spielen sie beim Kreditportfoliomanagement eine viel bedeutendere Rolle, da die Qualität des Portfolios durch die Zahl der Kreditausfälle bestimmt wird.

Bei der Modellierung der Zahl der Kreditausfälle kann das bereits oben angesprochene Selektionsproblem durch ein Hürdenmodell leicht berücksichtigt werden. Erst wenn ein Kreditnehmer nämlich die „Hürde“ genommen hat, überhaupt einen Kredit zu bekommen, kann der Kredit auch ausfallen. In einem aktuellen Beitrag verwenden Dionne et al. (1996) ein solches Zählmodellen-Hürdenmodell im Rahmen eines Kredit-Scoring Modells.

Eine ausgezeichnete Monographie über Zählmodellen bietet Winkelmann (1997). In einem Überblicksartikel beschreibt Pohlmeier (1994) Möglichkeiten und Vorteile der Schätzung von Zählmodellen für Paneldaten. Aktuelle Anwendungsbeispiele für Zählmodellen bieten Harhoff et al. (2000) sowie Licht und Zoz (1998).

7 *Resümee*

In diesem Aufsatz werden verschiedene mikroökonomische Methoden zur Evaluation von Kreditausfallrisiken dargestellt. Fragen wie „Was bestimmt das Kreditausfallrisiko?“ oder „Wie hoch ist die Wahrscheinlichkeit, mit der ein Kreditnehmer mit einem Set bestimmter Eigenschaften ausfällt?“ können mit binären Probit- oder Logitmodellen analysiert werden. Kann ein Kredit mehrere Risikozustände annehmen, z. B. keine/teilweise/vollständige Rückzahlung, so sind geordnete Logit/Probitmodelle geeignete Instrumente. Da Kreditakten i. d. R. in Form eines Paneldatensatzes vorliegen – ein individueller Kredit also über einen längeren Zeitraum beobachtet wird – werden auch Paneldatenmodelle für diese Modelle diskreter Entscheidungen besprochen.

Wenn es um Einschätzungen darüber geht, den „Hazard“ eines Kreditausfalls zu berechnen, dann sind Verweildauermodelle die adäquaten Analyseinstrumente. Als „Hazard“ wird die Wahrscheinlichkeit bezeichnet, mit der ein Kredit in Kürze ausfällt, gegeben dass er bis zu einem bestimmten Zeitpunkt noch nicht ausgefallen ist. Drei Verfahren stehen hierbei dem Analysten zur Verfügung. Der Kaplan-Meier-Schätzer dient dazu, erste Erkenntnisse über den Verlauf der „Überlebensfunktion“, also über die Wahrscheinlichkeit des Nicht-Ausfalls, zu gewinnen. Außerdem ist der Kaplan-Meier-Schätzer geeignet, Auskunft über sinnvolle Parametrisierungen von Modellen der „proportionalen Hazardrate“ zu geben, bei denen bei der Analyse der Überlebenswahrscheinlichkeit auch erklärende Variablen einbezogen werden.

Zu allen hier besprochenen Modellen werden sowohl Spezifikationstests als auch Gütemaße vorgestellt. Zusätzlich werden weiterführende Literaturhinweise

gegeben. Ebenso werden anhand eines aus Kreditakten zusammengestellten Datensatzes Beispiele und Interpretationshilfen gegeben.

Literaturverzeichnis

- Altman, E.I. und Saunders, A. (1998), Credit Risk Measurement: Developments over the last 20 years, *Journal of Banking & Finance*, Vol. 21, S. 1721-1742.
- Arulampalam, W. (1999), A Note on Estimated Coefficients in Random Effects Probit Models, *Oxford Bulletin of Economics and Statistics* 61(4), 597-602.
- Backhaus, K., B. Erichson, W. Plinke und R. Weiber (2000), *Multivariate Analysemethoden*, Heidelberg.
- Baltagi, B.H. (1995), *Econometrics of Panel Data*, Chichester.
- Basel Committee on Banking Supervision (1988), *International Convergence of Capital Measurement and Capital Standards*, Publication No.4.
- Basel Committee on Banking Supervision (1999), *A new Capital Adequacy Framework*, Publication No. 50.
- Basel Committee on Banking Supervision (2000a), *Range of Practice in Banks' Internal Ratings Systems*, Discussion Paper, Publication No. 66.
- Basel Committee on Banking Supervision (2000b), *Credit Ratings and Complementary Sources of Credit Quality Information*, Newsletter No. 3.
- Bester, H. (1985), Screening vs. Rationing in Credit Markets with Imperfect Information, in: *American Economic Review*, Vol. 75, pp. 850-859.
- Bester, H. (1987), The Role of Collateral in Credit Markets with Imperfect Information, in: *European Economic Review*, Vol. 31, pp. 887-899.
- Blossfeld, H.-P. und G. Rohwer (1995), *Techniques of Event History Modeling: New Approaches to Causal Analysis*, Mahwah
- Blossfeld, H.-P., A. Hamerle und K.-U. Mayer (1986), *Ereignisanalyse: Statistische Theorie und Anwendung in den Wirtschafts- und Sozialwissenschaften*, Frankfurt.
- Blundell, R., F. Laisney und M. Lechner (1993), Alternative Interpretations of Hours Information in an Econometric Model of Labour Supply, *Empirical Economics* 18, 393-415.
- Börsch-Supan, A.F. (1987), *Econometric Analysis of Discrete Choice – with Applications on the Demand for Housing in the U.S. and in West Germany*, Heidelberg.
- Boyes, W.J., D.L. Hoffman und S.A. Low (1989), An Econometric Analysis of the Bank Credit Scoring Problem, *Journal of Econometrics* 40, 3-14.
- Chamberlain, G. (1980), Analysis of Covariance with Qualitative Data, *Review of Economic Studies* 47, 225-238.
- Chesher, A., T. Lancaster und M. Irish (1985), On Detecting the Failure of Distributional Assumptions, *Annales de l'INSEE* 59/60, 7-45.
- Chesher, A. und M. Irish (1987), Residual Analysis in the Grouped and Censored Normal Linear Model, *Journal of Econometrics* 34, 33-61.
- Cox, D.R. (1972), Regression Models and life-tables, *Journal of the Royal Statistical Society* 34, 187-220.
- Cox, D.R. (1975), Partial Likelihood, *Biometrika* 62, 269-276.
- Davidson, R. und J. G. Mackinnon (1993), *Estimation and Inference in Econometrics*, Oxford University Press.
- Dionne, G., M. Artís und M. Guillen (1996), Count Data Models for A Credit Scoring System, *Journal of Empirical Finance* 3, 303-325.
- Dolton, P. und W. van der Klaaw (1999) The Turnover of Teachers: A Competing Risk Explanation, *The Review of Economics and Statistics* 81(3), 543-552.
- Eckey, H.-F., R. Kosfeld und C. Dreger (1995), *Ökonometrie: Grundlagen, Methoden, Beispiele*, Wiesbaden.
- Ebling, G. und N. Janz (1999), Export and Innovation Activities in the German Service Sector: Empirical Evidence at the Firm Level, ZEW discussion paper 99-53.

- Elsas, R., S. Henke, A. Machauer, R. Rott und G. Schenk (1997): Empirical analysis of credit relationships in small firms financing: Sampling design and descriptive statistics, Working Paper 98-06 (Center for Financial Studies Frankfurt).
- Estrella, A. (1998), A New Measure of Fit for Equations with Dichotomous Dependent Variables, *Journal of Business and Economics Statistics* 16(2), 198-205.
- Fahrmeir, L., W. Häußler und G. Tutz (1984), Diskriminanzanalyse, in: Fahrmeir, L. und A. Hamerle (Hrsg.), *Multivariate statistische Verfahren*, de Gruyter, 301-379.
- Gabler, S., F. Laisney und M. Lechner (1993), Semiparametric Estimation of Binary-Choice Models With an Application to Labor-Force Participation, *Journal of Business and Economic Statistics* 11(1), 61-80.
- Glewwe, P. (1997), A Test of the Normality Assumption in the Ordered Probit Model, *Econometric Reviews* 16(1), 1-19.
- Gouriéroux, C. (1989), *Économétrie des Variables Qualitatives*, Paris.
- Gouriéroux, C., A. Montfort und A. Trognon (1985), A General Approach to Serial Correlation, *Econometric Theory* 1, 315-340.
- Gouriéroux, C. und A. Montfort (1989), *Statistics and Econometric Models*, Cambridge.
- Gouriéroux, C., A. Montfort, E. Renault und A. Trognon (1987), Simulated Residuals, *Journal of Econometrics* 34, 201-252.
- Granger, C.W. und P. Newbold (1974), Spurious Regressions in Econometrics, *Journal of Econometrics* 2, 111-120.
- Greene, W.H. (1997), *Econometric Analysis*, Upper Saddle River.
- Hamerle, A. und G. Ronning (1995), Panel Analysis for Qualitative Variables, in: Arminger, G., C.C. Clogg und M.E. Sobel (Hrsg.), *Handbook of Statistical Modeling for the Social Sciences*, London, 401-450.
- Hamilton, J.D. (1994), *Time Series Analysis*, Princeton.
- Han, A. und J.A. Hausman (1990), Flexible Parametric Estimation of Duration and Competing Risk Models, *Journal of Applied Econometrics* 5, 1-28.
- Harhoff, D., F. Narin, F.M. Scherer und K. Vopel (2000), Citation Frequency and the Value of Patented Innovation, *Review of Economics and Statistics* 81(3), 511-515.
- Harhoff, D., K. Stahl und M. Woywode (1998), Legal Form, Growth and Exit of West German Firms: Empirical Results for Manufacturing, Construction, Trade and Service Industries, *Journal of Industrial Economics* 46(4), 453-488.
- Hausman, J.A. (1978), Specification Tests in Econometrics, *Econometrica* 46, 1251-1271.
- Hausman, J.A., A.W. Lo und A.C. MacKinlay (1992), An Ordered Probit Analysis of Transaction Price Changes, *Journal of Financial Economics* 31, 319-379.
- Jenkins, S.P. (1995), Easy Estimation for Discrete-Time Duration Models, *Oxford Bulletin of Economics and Statistics* 57(1), 129-137.
- Kaiser, U. (1997), Kaiser, U. (1997): The Determinants of BUND-future price changes: An ordered probit analysis using DTB and LIFFE data, ZEW Discussion paper No. 97-09
- Kaiser, U. (1999a), New Technologies and the Demand for Heterogeneous Labor: Firm-level Evidence for German Business-related Services, ZEW discussion paper 99-07, erscheint demnächst in: *Economics of Innovation and New Technology*.
- Kaiser, U. (1999b), Measuring Knowledge Spillovers in Manufacturing and Services: An Empirical Assessment of Alternative Approaches, ZEW discussion paper 99-62 (revised version).
- Kaiser, U. und F. Pfeiffer (2000), Collective Wage Agreements and the Adjustment of Workers and Hours in German Service Firms, ZEW discussion paper 00-33.
- Kalbfleisch, J.D. und R.L. Prentice (1980), *The Statistical Analysis of Failure Time Data*, New York.
- Kiefer, N.M. (1988), Economic Duration Data and Hazard Functions, *Journal of Economic Literature* 29(June), 646-679.
- King, G. und L. Zeng (1999), *Logistic Regression in Rare Events Data*, Harvard University mimeo.
- Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge.

- Laisney, F. und M. Lechner (1996), Combining Panel Data and Macro Information for the Estimation of a Panel Probit Model, *Zeitschrift für Wirtschafts- und Sozialwissenschaften* 116, 339-358.
- Laisney, F., M. Lechner und S. Strøm (1991), Lessons from Specification Tests for a Labour Supply Model, *Annales D'Économie et de Statistique* 20/21, 194-217.
- Leamer, E.E. (1978), *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, New York.
- Lechner, M. (1995), Some Specification Tests for Probit Models Estimated on Panel Data, *Journal of Business and Economic Statistics* 13(4), 475-488.
- Licht, G. und K. Zoz (1998), Patents and R&D: An Econometric Investigation Using Applications for German, European and US Patents by German Companies, *Annales D'Économie et de Statistique* 49/50, 329-360.
- Maddala, G.S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge.
- Manski, C.F. und S.R. Lerman (1977), The Estimation of Choice Probabilities from Choice-Based Samples, *Econometrica* 45, 1977-1988.
- Mátyás, L. (Hrsg., 1999), *Generalized Method of Moments Estimation*, Cambridge.
- Mátyás, L. und R. Sevestre (Hrsg., 1996), *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, Dordrecht.
- McFadden, D. (1974), The Measurement of Urban Travel Demand, *Journal of Public Economics* 3, 303-328.
- McKelvey, R. und W. Zavoina (1975), A Statistical Model for the Analysis of Ordinal Level Dependent Variables, *Journal of Mathematical Sociology* 4, 103-120.
- Nadiri, M.I. und S. Rosen (1969), Interrelated Factor Demands Functions, *American Economic Review*, 457-471.
- Nickell, P., W. Perraudin und S. Varotto (2000), Stability of Rating Transitions, *Journal of Banking and Finance* 24, 203-227.
- Nguyen Van, P., U. Kaiser und F. Laisney (2000), The Performance of German Firms in the Business-Related Services Sector, ZEW discussion paper 00-34.
- Pohlmeier, W. (1992), On the Simultaneity of Innovations and Market Structure, *Empirical Economics* 94(1), 71-83.
- Pohlmeier, W. und H. Entorf (1990), Employment, Innovation and Export Activity: Evidence from Firm-Level Data, in: J.P. Florens, M. Ivaldi, J.J. Laffont und F. Laisney (Hrsg.), *Microeconometrics: Surveys and Applications*, Oxford.
- Pohlmeier, W. (1994), Panelökonometrische Modelle für Zähldaten, *Schweizerische Zeitschrift für Volkswirtschaft und Statistik* 130(3), 553-574.
- Pudney, S. und J. Thomas (1995), Specification Tests for the Competing Risk Duration Model: An Application to Unemployment Duration and Sectoral Movement, *Oxford Bulletin of Economics and Statistics* 57(3), 323-346.
- Stiglitz, J. und A. Weiss (1981), Credit Rationing in Markets with Imperfect Information, in: *American Economic Review*, Vol. 71, pp. 393-410.
- Tutz, G. und W. Hennevogl (1996), Random Effects in Ordinal Regression Models, *Computational Statistics & Data Analysis* 22, 537-557.
- Urban, D. (1993), *Logit-Analyse: Statistische Verfahren zur Analyse von Modellen mit qualitativen Response-Variablen*, Stuttgart, Jena, New York.
- van den Berg, G.J. und B. van der Klaauw (1998), Combining Micro and Macro Unemployment Duration Data, University of Amsterdam mimeo.
- Veall, M.R. und K.F. Zimmermann (1992), Pseudo R²'s in the Ordinal Probit Model, *Journal of Mathematical Sociology* 16(4), 333-342.
- Verbeek, M. (2000), *A Guide to Modern Econometrics*, Chichester.
- Weiss, A.A. (1997), Specification Tests in Ordered Logit and Probit Models, *Econometric Reviews*, 16(4), 361-391.

Winkelmann, R. (1997), *Econometric Analysis of Count Data*, Berlin.