

**JOHANN WOLFGANG GOETHE-UNIVERSITÄT  
FRANKFURT AM MAIN**

**FACHBEREICH WIRTSCHAFTSWISSENSCHAFTEN**

**Ulrich Kaiser/Andrea Szczesny**

**Einfache ökonometrische Verfahren für die  
Kreditrisikomessung: Logit- und Probit-Modelle**

**No.61  
Dezember 2000**



**WORKING PAPER SERIES: FINANCE & ACCOUNTING**

**Ulrich Kaiser/Andrea Szczesny\*\***

**Einfache ökonometrische Verfahren für die  
Kreditrisikomessung: Logit- und Probit-Modelle\***

**No.61  
Dezember 2000**

**ISSN 1434-3401**

---

\* Dipl.-Volkswirt *Ulrich Kaiser*, Zentrum für Europäische Wirtschaftsforschung, Forschungsbereich Industrieökonomik und Internationale Unternehmensführung, L 7,1, 68161 Mannheim und Zentrum für Finanzen und Ökonometrie an der Universität Konstanz. Dipl.-Wirtsch.-Inf. *Andrea Szczesny*, Lehrstuhl für Betriebswirtschaftslehre, insb. Controlling, Johann Wolfgang Goethe-Universität Frankfurt, Mertonstr. 17, 60054 Frankfurt am Main.

\*\* Dieser Beitrag hat von hilfreichen Kommentaren von *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. Sämtliche hier vorgestellte Schätzungen wurden am Lehrstuhl für Betriebswirtschaftslehre, insbes. Controlling an der Universität Frankfurt am Main durchgeführt.

# **Einfache ökonometrische Verfahren für die Kreditrisikomessung: Logit- und Probit-Modelle**

Dezember 2000

## *Zusammenfassung*

Dieser Beitrag stellt verschiedene ökonometrische Methoden zur Bewertung und Berechnung von Kreditausfallrisiken vor und wendet diese auf einen Datensatz sechs deutscher Universalbanken an. Im Mittelpunkt stehen dabei Logit- und Probitmodelle, mit deren Hilfe die Ausfallwahrscheinlichkeit eines Kredites geschätzt werden kann. Dabei werden auch moderne Verfahren zur Analyse von Paneldaten besprochen. 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 probit and logit models which enable the credit analyst to quantify the default probability of an individual credit. Recent developments in the analysis of panel data are also introduced. 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.

Keywords: credit risk, default risk, probit and logit models, panel data

JELClassification: C23 C25 G21

## **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. Geradezu 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 ausschließlich statistischen Methoden. Zudem stellte der Ausschuss fest, dass es noch erheblich an Datenmaterial mangle. Bei der bankinternen Quantifizierung von Risiken, und erst recht bei der erforderlichen Validierung der Rating-Systeme durch die Bankenaufsicht werden allerdings ausreichende 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 den gesammelten Erfahrungswerten gewichtet und zu einem Kredit-Scoring verdichtet. Zur Bestimmung der Faktoren und Gewichte werden teilweise einfache statistische Ver-

fahren eingesetzt, die subjektive Einschätzung von Kreditanalysten spielt jedoch meist eine übergeordnete Rolle. 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, wie z.B. Wu (1965) zeigt. 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 ökonometrischer Methoden, 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, den Ingenieur- und Naturwissenschaften sowie der Soziologie. Das Ziel dieser Arbeit liegt darin, den Zugang zu den Methoden zu erleichtern, indem ihre wesentlichen Eigenschaften im Kontext der Risikomessung anschaulich beschrieben werden. 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- und 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- bzw. 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- bzw. Logitmodellen abgebildet werden.

Ein Analyst wird zunächst daran interessiert sein, die Robustheit seiner Kreditausfall-schätzungen zu überprüfen und einen Indikator für die Güte seines Modells zu finden. In den Abschnitten 3.3 und 3.4 stellen wir Spezifikationstests und Gütemaße vor.

Auf Erweiterungsmöglichkeiten der binären und geordneten Probit- und Logitmodelle wird in Abschnitt 3.5 eingegangen. Dabei werden Paneldatenmodelle sowie Mehrgleichungsmodelle mit Simultanität und mit Korrelation der Fehlerterme skizziert. Abschnitt 0 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.<sup>1</sup> Die in diesem 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).<sup>2</sup> Um Verzerrungen der Stichprobe hinsichtlich problembehafteter bzw. nicht problembehafteter Kredite zu vermeiden, wurde der Datensatz hinsichtlich dieses Kriteriums geschichtet.<sup>3</sup> 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 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 Unternehmen, das Verwerten von Sicherheiten, die Fälligestellung von Krediten, Abwicklungen, Vergleiche, Konkurse und Sanierungen. Selbstverständlich können ökonometrische Ansätze sämtliche verschiedene Riskozustände modellieren. Aus Gründen der Übersichtlichkeit

---

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

2 Weitere auf dem Datensatz aufbauende empirische Studien finden sich auf der Homepage des CFS ([http://www.ifk-cfs.de/pages/veroef/cfswor/index\\_d.htm](http://www.ifk-cfs.de/pages/veroef/cfswor/index_d.htm)).

3 Auf Probleme von Stichprobenselektionsverzerrungen und deren Korrektur gehen wir in Abschnitt 3.5.2 ein.

und der erleichterten Zugangs zu den verschiedenen Schätzansätzen wird jedoch die Zusammenfassung in lediglich zwei bzw. drei Risikozustände vorgenommen.

*Kein Ausfall, Teilausfall und Vollaussfall:*

Wenn eine Unterteilung der Daten in die drei Kategorien kein Ausfall, Teilausfall und Vollaussfall 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 Vollaussfall münden. Dazu gehören Probleme, die beispielsweise zu einer Stundung von Zins- und Tilgungszahlungen, zum Einfordern zusätzlicher Sicherheiten oder zur Verwertung von Sicherheiten führen. Als Vollaussfall bezeichnen wir schwere Probleme bei der Erfüllung des Kreditvertrages, die zum Beispiel eine Fälligkeitstellung nach sich ziehen sowie Abwicklungen, Vergleiche und Konkurse. Aber auch Sanierungen, die im Erfolgsfall natürlich keinen Vollaussfall des Kredites bedeuten, werden zu der Kategorie „Vollaussfall“ 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. Einen guten Einstieg in die Literatur zu möglichen Indikatoren für das Kreditrisiko und dessen Modellierung bieten Altman und Saunders (1998) sowie die Veröffentlichungen des Basler Ausschusses für Bankenaufsicht (2000a und 2000b). Die folgenden Variablen werden für die empirische Analyse 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(Umsatz)`: Variable, welche die Größe des Unternehmens anhand der Höhe seiner Umsätze repräsentiert. Dabei werden die Umsätze mit Hilfe des natürlichen Logarithmus transformiert.
- `ln(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 langfristigen Passiva und mittel- und langfristigen 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.<sup>4</sup> 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. 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  $\mathbf{e}_i$  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:

---

4 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.

$$Ausfall_i = \begin{cases} 1 & \text{wenn } Ausfall_i^* = x_i \mathbf{b} + \mathbf{e}_i > s \\ 0 & \text{sonst.} \end{cases} \quad (1)$$

Sofern der Fehlerterm  $\mathbf{e}_i$  einer symmetrischen Verteilung folgt, kann Gleichung (1) in folgenden Wahrscheinlichkeiten ausgedrückt werden:

$$\begin{aligned} P(Ausfall_i = 1 | x_i) &= P\left(\frac{\mathbf{e}_i > s - x_i \mathbf{b}}{\mathbf{s}}\right) = 1 - F\left(\frac{\mathbf{e}_i \leq s - x_i \mathbf{b}}{\mathbf{s}}\right) \\ P(Ausfall_i = 0 | x_i) &= P\left(\frac{\mathbf{e}_i \leq s - x_i \mathbf{b}}{\mathbf{s}}\right) = F\left(\frac{\mathbf{e}_i \leq s - x_i \mathbf{b}}{\mathbf{s}}\right), \end{aligned} \quad (2)$$

wobei  $F$  im Logitfall die Verteilungsfunktion der logistischen Verteilung und im Probitfall die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Der Parameter  $\sigma$  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  $\mathbf{s} = \Pi/\sqrt{3}$  und  $s=0$  für das Logitmodell sowie  $\mathbf{s} = 1$  und  $s=0$  für das Probitmodell. Diese Normierungen ändern an der Interpretierbarkeit des Modells nichts.<sup>5</sup> Soweit es im Text nicht anders vermerkt ist, beziehen sich die nun folgenden Ausführungen auf den Probitfall.

Die Berechnung des Koeffizientenvektors  $\mathbf{b}$  erfolgt mit der Maximum-Likelihood-Methode. Dabei geht es darum, über die Wahl von  $\mathbf{b}$  die Wahrscheinlichkeit zu maximieren, mit der der vorliegende Datensatz reproduziert wird.<sup>6</sup>

Die Schätzung von Paneldatenmodellen, die in Abschnitt 3.4.1 beschrieben wird, verlangt die Equidistanz, also einen zeitlich gleichen Abstand zwischen den einzelnen Beobachtungszeitpunkten. Dies ist im uns vorliegenden Datensatz nicht der Fall. Vielmehr existieren zu einigen Krediten mehr als eine Beobachtung innerhalb eines Jahres. Die

---

5 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  $\sigma$ ).

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.

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

---

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

signifikant größer sind als die der Vorjahre. Die konjunkturelle Lage hat also einen bedeutenden Effekt auf die Kreditausfallwahrscheinlichkeit.

Tabelle 1: Ergebnisse des Probitmodells

<b>Variable</b>	<b>Koeff.</b>	<b>Std.fehler</b>	<b>p-Wert</b>
Ln(Umsatz)	-0,8744	0,8333	0,294
Ln(Umsatz) <sup>2</sup>	0,0370	0,0361	0,306
Eigenkapitalquote	-2,8844	0,4011	0,000
Cash-Flow	-0,4113	0,1735	0,018
Anlagedeckungsgrad	-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
<b>Wald-Tests auf gemeinsame Signifikanz</b>			
<b>Variablen</b>	<b>chi<sup>2</sup></b>	<b>d.o.f</b>	<b>p-Wert</b>
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

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.

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 oder Semi-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 einer 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: 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. Sie gibt an, 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-/Logitmodellen 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  $\mathbf{s}=1$ , gegeben durch:

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

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  $-x_i \mathbf{b}$  abhängt. Dennoch lässt sich aus dem Vorzeichen des Koeffizienten  $\mathbf{b}_k$  der Effekt auf die Wahrscheinlichkeitsveränderung eindeutig ablesen: Ein positives (negatives) Vorzeichen von  $\mathbf{b}_k$ , 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:

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

wobei  $x_l$  eine Dummy-Variable bezeichnet. Ebenso wie für den Fall kontinuierlicher Variablen gibt das Vorzeichen des Koeffizienten  $\beta_l$  Auskunft über den Effekt der erklärenden Dummy-Variable  $x_l$ . 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.

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.

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.

Tabelle 2: Marginale Effekte der Probitschätzung

<b>Variable</b>	<b>Koeff.</b>	<b>Std.fehler</b>	<b>p-Wert</b>
ln(Umsatz)	-0,2640	0,2515	0,294
ln(Umsatz) <sup>2</sup>	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

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:

$$\begin{aligned}
 P(\text{Ausfall}_i = 1) &= 1 - F[-\{\ln(220\,000) \cdot \beta_{\ln(\text{Umsatz})} \\
 &+ \ln(220\,000)^2 \cdot \beta_{\ln(\text{Umsatz})^2} + 0,2 \cdot \beta_{\text{Eigenkapitalquote}} + 0,58 \cdot \beta_{\text{Cash-Flow}} \\
 &+ 0,8 \cdot \beta_{\text{Anlagendeckung}} + 0 \cdot \beta_{\text{bhaft}} + \beta_{1995} + \beta_{\text{Handel}} + \beta_{\text{Konstante}}\}] \quad (5) \\
 &= 1 - F[-\{12,3 \cdot (-0,87) + 151,3 \cdot 0,04 + 0,2 \cdot (-2,88) + 0,58 \cdot (-0,41) \\
 &+ 0,8 \cdot (-0,13) + 0,67 - 0,13 + 4,69\}] = 1 - F[0,3368] = 0,37
 \end{aligned}$$

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,<sup>7</sup> 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 von größer als 1 und für 66 Kreditnehmer eine Kreditausfallwahrscheinlichkeit von kleiner als 0 ergeben.

---

7 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.



### 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:

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

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, identisch und unabhängig verteilten Störterm beschrieben. Auch hier Restriktionen auferlegt werden. STATA wählt die Restriktion  $\sigma = 1$  und setzt den Wert der Konstanten auf 0; es werden alle Schwellenwerte  $s$  mitgeschätzt, Limdep normiert  $\sigma$  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:

$$\begin{aligned}
\frac{\partial P(\text{Ausfall}_i = 2 | x_i)}{\partial x_{ik}} &= f(s_2 - x_i \mathbf{b}) \mathbf{b}_k \\
\frac{\partial P(\text{Ausfall}_i = 1 | x_i)}{\partial x_{ik}} &= (f(s_1 - x_i \mathbf{b}) - f(s_2 - x_i \mathbf{b})) \mathbf{b}_k, \\
\frac{\partial P(\text{Ausfall}_i = 0 | x_i)}{\partial x_{ik}} &= -f(s_1 - x_i \mathbf{b}) \mathbf{b}_k.
\end{aligned} \tag{7}$$

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 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.<sup>8</sup> Die Vorzeichen der Koeffizienten im geordneten Probitmodell geben die Richtung an, mit denen sich die Wahrscheinlichkeiten der Randkategorien verändern. In diesem Fall sind die Randkategorien die Wahrscheinlichkeit, mit der keine Probleme auftreten, und die Wahrscheinlichkeit, mit der erhebliche Probleme auftreten. Über die dazwischen liegende Kategorie, d. h. die Wahrscheinlichkeit, mit der es zu einem Teilausfall kommt, kann anhand dieser Tabelle keine Auskunft gegeben werden.

Tabelle 3: Ergebnisse des geordneten Probitmodells

<b>Variable</b>	<b>Koeff.</b>	<b>Std.fehler</b>	<b>p-Wert</b>
ln(Umsatz)	-0,7219	0,7946	0,364
ln(Umsatz) <sup>2</sup>	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 <sub>1</sub>	-3,4718	4,5400	0,444
s <sub>2</sub>	-2,9170	4,5397	0,521
<b>Wald-Tests auf gemeinsame Signifikanz</b>			
<b>Variablen</b>	<b>chi<sup>2</sup></b>	<b>d.o.f</b>	<b>p-Wert</b>
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 0,8 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

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

einer Verteilungsfunktion immer unverändert, nämlich eins, bleibt. 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) <sup>2</sup>	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. Standardfehler für die marginalen Effekte können mit der „ $\delta$ -Methode“ (Z.B. Greene 1997, Kap. 6.7.5) berechnet werden.

An dieser Stelle wird der Vorteil eines geordneten Probit-Modells gegenüber dem binären Probitmodell deutlich. Der Kreditbeurteiler erhält mehr Informationen 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.<sup>9</sup>

Geordnete Logit- und Probitmodelle können für die Berechnung von Übergangswahrscheinlichkeiten in verschiedene Riskozustände berechnet werden. So verwenden Nickell et al. (2000) geordnete Probitmodelle zur Schätzung solcher Markovketten.<sup>10</sup> Ein

<sup>9</sup> 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, würden im vorliegenden Beispiel die Ergebnisse der Spezifikation ohne die quadrierten Umsätze verwendet.

<sup>10</sup> Kaiser und Szczesny (2000) gehen auf Markovkettenmodelle im Zusammenhang mit der Kreditrisikomessung genauer ein.

Nachteil des Vorgehens von Nickell et al. (2000) ist es jedoch, dass die qualitative Effekte erklärender Variablen auf die Übergänge in die verschiedenen Riskozustände per Konstruktion für alle Arten der Übergänge gleich sind. So ist beim geordneten Probitmodell der qualitative Effekt z.B. des Cash Flows auf den Übergang von „kein Ausfall“ zu „teilweiser Ausfall“ der gleiche wie auf den Übergang von „kein Ausfall“ zu „vollständiger Ausfall“. Insofern erscheint es wünschenswert, mögliche differenziellen Effekt der erklärenden Variablen auf verschiedene Riskozustände zuzulassen. Ein solches Vorgehen ermöglicht das multinomiale Logitmodell (s. Abschnitt 3.5.4), das von Kaiser (2000a) für die Analyse von Zahlungsschwierigkeiten im Rahmen eines Markovkettenmodells verwendet wird.

### 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 vorhersagen kann.

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 so genannte 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 für eine Fehlspezifikation des Modells. Aus mehreren weiteren Gründen ist die hohe Bedeutung, die 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 in den siebziger Jahren und Anfang der achtziger Jahre häufig ein  $R^2$  von 0,99 aus, z. B. der wichtige Beitrag von Nadiri und Rosen (1969) zu dynamischen Faktor-

nachfragesystemen. 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 linearer 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  $R^2$  sind leicht manipulierbar.<sup>11</sup>

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.<sup>12</sup>

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 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.<sup>13</sup> 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  $\mathbf{b}$  ge-

---

<sup>11</sup> Leamer (1978, Kap. 3) nimmt dazu ausführlich Stellung.

<sup>12</sup> Ausnahmen sind z. B. extrem volatile Finanzmarktdaten, z. B. Intradagesdaten (s. die Sonderausgabe des *Journal of Business and Economic Statistics* 18 (2), April 2000).

geschätzt wird, sondern das Verhältnis  $\mathbf{b}/\mathbf{s}$ . Durch die Normierung  $\mathbf{s} = 1$  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  $\mathbf{s}_i = h(\mathbf{g} z_i)$ , wobei  $\mathbf{g}$  einen Koeffizientenvektor bezeichnet und  $z_i$  diejenigen Variablen bezeichnet, die Heteroskedastie hervorrufen, so wird anstatt des Parametervektors  $\mathbf{b}$  das Verhältnis  $\mathbf{b}/\exp(z_i \mathbf{g})$  geschätzt.<sup>14</sup> 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.<sup>15</sup> Weil der Wert der latenten Variablen, in diesem Fall der Wert von  $\text{Ausfall}_i$ , 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:

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

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

$$\mathbf{z}_i^g = \frac{f(s_{j-1} - x_i \mathbf{b}) - f(s_j - x_i \mathbf{b})}{F(s_j - x_i \mathbf{b}) - F(s_{j-1} - x_i \mathbf{b})}, \quad (9)$$

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 welcher starken Abweichungen es zwischen den Koeffizientenvektoren  $\beta$  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 (Gouriéroux 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  $\mathbf{g}=0$ , 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,  $\mathbf{e}_i^g x_i$ , und (ii) der Interaktion zwischen dem generalisierten Residuum und sowohl dem gefitteten Wert der Ursprungsschätzung,  $x_i \mathbf{b}$ , als auch den Variablen, die im Verdacht stehen, Heteroskedastizität hervorzurufen,  $(\mathbf{e}_i^g x_i \mathbf{b})z_i$ , auf einen Vektor von Einsen ohne den Einschluss einer Konstanten.<sup>16</sup> Analoges gilt für die generalisierten Residuen der geordneten Modelle. Die sich als  $N \cdot R^2$  ergebende Teststatistik ist  $\chi^2$  verteilt mit Anzahl der Freiheitsgrade gleich Anzahl der im Verdacht Heteroskedastie hervorzurufen stehenden Variablen  $z$ .<sup>17</sup>

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  $\mathbf{e}_i^g x_i$  und  $(\mathbf{e}_i^g x_i \mathbf{b})x_i$  auf einen Einervektor. 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  $\chi^2$ -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

---

16 Es handelt sich hierbei um eine so genannte „künstliche Regression“, Davidson und MacKinnon (1993, Kap. 6) gehen näher auf dieses in der Ökonometrie sehr wichtige Verfahren ein. Die Terme  $\mathbf{e}_i^g x_i$  bzw.  $(\mathbf{e}_i^g x_i \mathbf{b})z_i$  bezeichnen die Scores (1. Ableitungen) der Log-Likelihoodfunktion nach  $\beta$  bzw.  $\gamma$ .

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. Gouriéroux et al. (1985) schlagen einen Test auf Autokorrelation bei binären und geordneten Auswahlmodellen vor.



Probitmodell kann Homoskedastizität auf den konventionellen Signifikanzniveaus nicht verworfen werden.

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

$$P(\text{Ausfall}_i = 1 | x_i) = 1 - F\left(\frac{\mathbf{e}_i \leq s - x_i \mathbf{b}}{\mathbf{s}}\right) \quad (2)$$

würde dann

$$P(\text{Ausfall}_i = 1 | x_i) = 1 - F\left(\frac{\mathbf{e}_i \leq s - x_i \mathbf{b}}{h(z_i \mathbf{g})}\right) \quad (2')$$

geschätzt. Standardsoftwareprogramme wie STATA und Limdep spezifizieren die Funktion  $h(\cdot)$  als Exponentialfunktion. Analoges gilt für die geordneten Modelle.<sup>18</sup> 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  $\mathbf{e}_i^g x_i$ ,  $\mathbf{e}_i^g (x_i \mathbf{b})^2$  und  $(\mathbf{e}_i^g x_i \mathbf{b})^3$  auf einen Einervektor 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  $\chi^2$ -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  $\chi^2$ -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 semiparametri-

schen 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.<sup>19</sup> Informationsmatrixtests auf Normalität im geordneten Wahrscheinlichkeitsmodellen werden von Glewwe (1997) und Weiss (1997) vorgeschlagen. Einen Überblick über Informationsmatrixtests bieten Gouriéroux und Montfort (1995, Kap. 18.5).

### 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 so genannte „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  $x_i \hat{\mathbf{b}}$ ) den Wert 0 übersteigt und umgekehrt:

$$\begin{aligned} \hat{Ausfall}_i^* &= 1 \text{ wenn } x_i \hat{\mathbf{b}} > 0 \\ \hat{Ausfall}_i^* &= 0 \text{ wenn } x_i \hat{\mathbf{b}} \leq 0 \end{aligned} \tag{10}$$

---

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.

wobei  $\hat{Ausfall}_i^*$  und  $x_i \hat{\mathbf{b}}$  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.<sup>20</sup> Wenn ein Zustand seltener als in fünf Prozent der Fälle eintritt, so müssen einzelne Zustände entweder zusammengefasst oder ggf. so genannte „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.<sup>21</sup>

Mit den in Tabelle 1 abgebildeten Ergebnissen korrespondiert die „prediction/realization“-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 vergebener Kredit kann ausfallen, ein fälschlicherweise abgelehnter guter Kreditkunde 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.<sup>22</sup>

---

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.

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

Tabelle 5: Trefferquoten des binären Probitmodells

		vorhergesagt		
		Ausfall: nein	Ausfall: ja	
tatsächlich	Ausfall: nein	645 94,02%	41 5,98%	686
	Ausfall: ja	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 einem oder mehr Schwellenwerten besteht. So geben Standardsoftwareprogramme häufig das Bestimmtheitsmaß von McFadden (1974) an, das sich wie folgt berechnet:

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

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: 116) unter Berufung auf Urban (1993: 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 bestimmt. McFad-

dens Pseudo-R<sup>2</sup> und das R<sup>2</sup> der linearen Regression haben nichts miteinander gemein. Vor diesem Hintergrund ist es auch unverständlich, dass Eckey et al. (1995: 179) McFaddens Pseudo-R<sup>2</sup> als „Unsicherheitsmaß“ bezeichnen. Aufgrund einer impliziten Bestrafung hoher Stichprobengrößen liegt das Pseudo-R<sup>2</sup> bei kleinen Stichproben generell höher als bei großen Stichproben.

Für das binäre Probitmodell aus Tabelle 1 ergibt sich ein McFadden-Pseudo-R<sup>2</sup> 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-R<sup>2</sup> 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-R<sup>2</sup>.

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:

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

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-R<sup>2</sup> 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 laten-

ten 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.<sup>23</sup> Dieses Gütemaß berechnet sich als

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

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  $-2*[\ln(L')-\ln(L'')]$  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  $\chi^2$ -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  $\chi^2$ -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.

---

<sup>23</sup> Estrella (1998) berücksichtigt dabei allerdings nicht das Pseudo- $R^2$  von McKelvey und Zavoina (1975).

### 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 wurden – 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 in der Regel 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. Beispielsweise 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:

$$Ausfall_i = \begin{cases} 1 & \text{wenn } Ausfall_i^* = x_{it} \mathbf{b} + \mathbf{e}_{it} = x_{it} \mathbf{b} + \mathbf{a}_i + u_{it} \geq s \\ 0 & \text{sonst.} \end{cases}, \quad (1')$$

wobei der Vektor der erklärenden Variablen  $x_{it}$  keine Konstante enthält und  $\mathbf{a}_i$  einen individuenspezifischen Effekt repräsentiert. Dieser individuenspezifische Effekt wird als über die Zeit konstant angenommen und häufig als „unbeobachtbare Heterogenität“ bezeichnet. Wenn  $\mathbf{a}_i$  über alle Individuen  $i$  konstant ist und Unabhängigkeit der Fehler-

terme  $u_{it}$  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 individualspezifischen Konstanten  $\mathbf{a}_i$  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  $\mathbf{a}$  – 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 Kreditrisikomessung von hoher Bedeutung sind, verzichten wir auf die weitere Darstellung des Fixed-Effects-Schätzers.<sup>24</sup>

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

24 Für das Probitmodell existiert bislang nur der Fixed-Effects-Schätzansatz. Für das Logitmodell existieren sowohl Fixed- als auch Random-Effects-Schätzer vorhanden. Daher konzentriert sich die Diskussion von binären Paneldatenmodellen in diesem Abschnitt auf Logitmodelle.

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



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  $\mathbf{a}_i$ 's) und den Residuen (den  $u_{it}$ 's) bzw. den erklärenden Variablen und den Individualeffekten. 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).<sup>26</sup> Die Teststatistik lautet wie folgt:

$$\mathbf{c}^2 = (\mathbf{b}_{RE} - \mathbf{b}_{FE})' (\mathbf{\Omega}_{FE} - \mathbf{\Omega}_{RE})^{-1} (\mathbf{b}_{RE} - \mathbf{b}_{FE}), \quad (14)$$

wobei  $\mathbf{b}_{RE}$  und  $\mathbf{b}_{FE}$  ( $\mathbf{\Omega}_{RE}$  und  $\mathbf{\Omega}_{FE}$ ) die geschätzten Parametervektoren (Varianz-Kovarianzmatrizen) des Random- bzw. des Fixed-Effects-Modells bezeichnen. Diese Teststatistik ist  $\mathbf{c}^2$ -verteilt mit Anzahl der Freiheitsgrade gleich dem Rang der Matrix  $(\mathbf{\Omega}_{FE} - \mathbf{\Omega}_{RE})$ .<sup>27</sup> Die Nullhypothese lautet, dass das Random-Effects-Modell korrekt ist, 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  $\mathbf{c}^2$ -Verteilung, so kann das Random-Effects-Modell nicht abgelehnt werden.

---

<sup>26</sup> Eine ausführliche Darstellung der Familie der Hausman-Tests geben Gouriéroux und Montfort (1995, Kap. 18.4).

<sup>27</sup> 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

<b>Variable</b>	<b>Koeff.</b>	<b>Std.fehler</b>	<b>p-Wert</b>
ln(Umsatz)	-4,4770	2,9888	0,134
ln(Umsatz) <sup>2</sup>	0,1924	0,1302	0,139
Eigenkapitalquote	-7,2281	1,3809	0,000
Cash-Flow	-1,1850	0,5195	0,023
Anlagedeckungsgrad	-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	24,3538	16,9114	0,150
Corr.	2,2424	0,2996	0,000
<b>Wald-Tests auf gemeinsame Signifikanz</b>			
<b>Variablen</b>	<b>chi<sup>2</sup></b>	<b>d.o.f</b>	<b>p-Wert</b>
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 Probitchätzungen 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 Paneldatensatzes zulässig ist, also ob die Panelstruktur irrelevant ist oder ob ein einfaches binäres Logitmodell geschätzt werden sollte. Die korrespondierende Hausman-Teststatistik lautet:

$$\mathbf{c}^2 = (\mathbf{b}_{FE} - \mathbf{b})'(\boldsymbol{\Omega}_{FE} - \boldsymbol{\Omega})^{-1}(\mathbf{b}_{FE} - \mathbf{b}), \quad (15)$$

wobei  $\mathbf{b}$  und  $\boldsymbol{\Omega}$  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  $\chi^2$ -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 Hennevogl (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.<sup>28</sup> Ein wesentlicher Beitrag jenes Artikels ist es, dass gezeigt wird, dass die Verwendung von Makrovariablen als zusätzliche Momentenbedingungen zu deutlichen Effizienzgewinnen führt.

---

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

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-R<sup>2</sup>-Gütemaße verwendet werden.

### 3.5.2 *Multivariate Probitmodelle*

Beim multivariaten Probitmodell wird davon ausgegangen, dass die Fehlerterme von zwei (bivariaten) oder mehr als zwei (multivariaten) Probitgleichungen miteinander korreliert sind.<sup>29</sup> Eine getrennte Schätzung würde dann zu konsistenten, aber ineffizienten Schätzern führen.<sup>30</sup> Multivariate Probitmodelle tragen 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 Probitmodellen 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.<sup>31</sup> 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.

---

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).

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

Stichprobenselektionsverzerrungen kann man auch mit Hilfe des „choice-based“-Schätzers von Manski und Lerman (1977) in den Griff bekommen. 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 Kreditscoringproblem bei Kreditkartenbesitzern an. Da der Kreditanalytist 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 vorhandenen 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 Selektionskorrektur oder einen „choice-based sampling“-Ansatz angewendet werden. Dies trifft natürlich für fast alle Arbeiten zum Kreditrisikoausfallrisiko zu, da zum einen Daten über abgelehnte Kredite selten gesammelt werden und zum anderen eine Grundgesamtheit nur schwer zu bestimmen ist.

### 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 gegenseitig 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 findet sich in Maddala (1983), weitere Anwendungen und Verallgemeinerungen geben Pohlmeier und Entorf (1990) sowie Pohlmeier (1992).

#### 3.5.4 *Multinomiale Logitmodelle*

Bislang wurden an diesem Aufsatz lediglich Modelle für geordnete Variablen untersucht. Folgt die zu untersuchende Variable jedoch keinem natürlich geordneten Zustand, dann stellen geordnete Logit- und Probitmodelle nicht mehr die geeignete Analyseform dar. Soll z.B. untersucht werden, für welche Hausbank sich ein Kreditnehmer entscheidet, so handelt es sich um ein Entscheidungsproblem, dass – zumindest aus Sicht des Kreditnehmers – ökonometrisch gesehen „ungeordnet“ ist. Im Kern handelt es sich beim multinomialen Logitmodell um eine Erweiterung des binären Logitmodells eben für mehrere ungeordnete Zustände.

Multinomiale Logitmodelle können auch zur Berechnung von Markovketten verwendet werden. Kaiser (2000a) und Nguyen van et al. (2000) wenden multinomiale Logitmodelle an. Beide Autoren führen zudem random effects in die Schätzungen, Nguyen van et al. (2000) erweitern das multinomiale Logitmodelle ausserdem zu einem „generalized additive model“ (Hastie und Tibshirani 1987), also um einen nichtparametrisch geschätzten Effekt einer erklärenden Variablen.

Eine wenig wünschenswerte Eigenschaft multinomialer Logitmodelle ist die „Unabhängigkeit von irrelevanten Alternativen“, die besagt, dass das Verhältnis zweier Auswahlwahrscheinlichkeiten unabhängig von der Anzahl der Auswahlmöglichkeiten ist (z.B. Greene 1997, Kap. 19.7.3). Hausman und McFadden (1984) haben einen Test auf Gültigkeit der Unabhängigkeitsannahme vorgeschlagen. Kann Unabhängigkeit nicht ge-

nommen werden, so stellen genistete Logitmodelle (z.B. Kaiser 2000b) und multivariate Probitmodelle (Lerman und Manski 1981) Alternativen dar.

#### **4 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 in der Regel in Form eines Paneldatensatzes vorliegen – ein individueller Kredit also über einen längeren Zeitraum beobachtet wird –, werden auch Paneldatenmodelle für diese Modelle diskreter Entscheidungen besprochen.

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.

## 5 Literaturverzeichnis

- Altman, E. I. und A. Saunders, (1998), Credit Risk Measurement: Developments over the Last 20 Years, *Journal of Banking & Finance* 21, 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, *American Economic Review* 75, 850-859.
- Bester, H. (1987), The Role of Collateral in Credit Markets with Imperfect Information, *European Economic Review* 31, 887-899.
- 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. und M. Irish (1987), Residual Analysis in the Grouped and Censored Normal Linear Model, *Journal of Econometrics* 34, 33-61.
- Davidson, R. und J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, Oxford.
- 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 No. 99-53, Mannheim.
- 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*, Berlin, 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.
- Han, A. und J. A. Hausman (1990), Flexible Parametric Estimation of Duration and Competing Risk Models, *Journal of Applied Econometrics* 5, 1-28.
- Hastie, T. und R. Tibshirani (1987), Generalized Additive Models: Some Applications. *Journal of the American Statistical Association*, 82(398), 371-86.
- 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.
- 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, Mannheim.
- Kaiser, U. (1999a), *New Technologies and the Demand for Heterogeneous Labor: Firm-Level Evidence for German Business-Related Services*, ZEW Discussion Paper No. 99-07, Mannheim, 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 No. 99-62, Mannheim (revidierte Fassung). erscheint demnächst in: *Research Policy*.
- Kaiser, U. (2000a), *Moving in and out of Financial Distress: Evidence for newly founded services sector firms*, ZEW mimeo, Mannheim.
- Kaiser, U. (2000b), *Research cooperation and research expenditures with endogeneous absorptive capacity: Theory and microeconomic evidence for the German service sector*, ZEW Discussion paper 00-25, Mannheim.
- Kaiser, U. und F. Pfeiffer (2000), *Collective Wage Agreements and the Adjustment of Workers and Hours in German Service Firms*, ZEW Discussion Paper No. 00-33, Mannheim.
- Kaiser, U. und A. Szczesny (2000), *Einfache ökonomische Verfahren für die Kreditrisikomessung: Verweildauermodelle*, Universität Frankfurt, Working Paper Series: Finance and Accounting, No. ??, Frankfurt.
- King, G. und L. Zeng (1999), *Logistic Regression in Rare Events Data*, Harvard University mimeo.
- 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.
- 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., Perraudin, W., Varotto, S. (2000), Stability of rating transition. *Journal of Banking & Finance* 24, 203--227.
- 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.
- Stiglitz, J. und A. Weiss (1981), Credit Rationing in Markets with Imperfect Information, *American Economic Review* 71, 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.
- Veall, M. R. und K. F. Zimmermann (1992), Pseudo R<sup>2</sup>'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. (1997), Specification Tests in Ordered Logit and Probit Models, *Econometric Review* 16 (4), 361-391.
- Wu, H.-K. (1969), Bank examiner criticisms, bank loan defaults, and bank loan quality. *Journal of Finance* 24, 697-705.