Quantitative Analyse dynamischer nichtlinearer Panelmodelle

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Agrarwissenschaften
der Georg-August-Universität Göttingen

vorgelegt von

Oliver Bode
geboren in Osterode am Harz

Göttingen, im November 2002

D 7

1. Referent:	Prof. Dr. H. Lauenstein
2. Korreferent:	Prof. Dr. M. Leserer

Tag der mündlichen Prüfung: 6. Juli 2001

Vorwort

Die vorliegende Arbeit wurde am Institut für Agrarökonomie der Georg-August-Universität Göttingen erstellt. An dieser Stelle möchte ich all den Personen Dank sagen, die in vielen Gesprächen und Diskussionen zum Entstehen dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt Herrn Prof. Dr. H. Lauenstein für die freundliche Überlassung des Themas und die wissenschaftliche Betreuung der Arbeit. Für die stets wertvollen Anregungen bei der Konzeption und Ausführung der Arbeit bin ich ihm besonders dankbar. Ferner gilt mein Dank Herrn Prof. Dr. M. Leserer für die Übernahme des Korreferats und den zu jeder Zeit gegebenen wertvollen fachlichen Rat.

Dank sagen möchte ich auch Frau D. Paulik-Krendel für die freundliche Hilfe bei der Lösung zahlreicher computertechnischer Probleme. Weiterhin möchte ich Frau M. Scholz und Frau D. Paulik-Krendel für die Ermutigungen und die Unterstützung bei der Erstellung der vorliegenden Arbeit Dank sagen.

Schließlich möchte ich meinen Eltern und meiner Frau Tina danken, die mir auf meinem Weg immer zur Seite standen. Tina ist diese Arbeit gewidmet.

Wiesbaden, im November 2002

Oliver Bode

INHALTSVERZEICHNIS

1 EINLEITUNG 1

1.1 Problemstellung 1

1.2 Zielsetzung und Vorgehensweise 2

2 STATISCHE NICHTLINEARE PANELMODELLE 7

2.1 Einführende Bemerkungen zur Verwendung von Paneldaten 7

2.2 Grundlagen der Spezifikation statischer nichtlinearer Panelmodelle 9

2.2.1 Statische nichtlineare Panelmodelle mit festen Effekten 13

2.2.2 Statische nichtlineare Panelmodelle mit zufälligen Effekten 16

2.2.2.1 Probitmodelle mit äquikorrelierten Störgrößen 16

2.2.2.2 Das statische Probitmodell für korrelierte α_i und x_{it} - Der Chamberlain-Ansatz 20

2.3 ML-Schätzung in binomialen Antwortmodellen 22

2.3.1 Grundzüge der Newton-Raphson-Methode 23

2.3.2 Schätzung der asymptotischen Kovarianzmatrix des ML-Schätzers 26

2.3.3 Zur Verwendung des Likelihood-Verhältnis-Tests 27

3 DYNAMISCHE NICHTLINEARE PANELMODELLE 29

3.1 Vorbemerkungen 29

3.2 Spezifikation dynamischer Probitmodelle 30

3.2.1 Statische Probitmodelle mit autokorrelierten Störgrößen 31

3.2.2 Dynamische Probitmodelle mit äquikorrelierten Störgrößen 33

3.2.3 Dynamische Probitmodelle mit autokorrelierten Störgrößen 37

4 GRUNDLAGEN DER IMPLEMENTIERUNG STATISCHER UND DYNAMISCHER PROBITMODELLE 39

4.1 Schätzung statischer Probitmodelle mit autokorrelierten Störgrößen - Der MECOSA-Ansatz 39

4.1.1 Modellspezifikation des statischen Probitmodells mit autokorrelierten Störgrößen in der Nomenklatur des MECOSA-Ansatzes 40

4.1.2 Parameterschätzung mittels des MECOSA-Algorithmus 41

4.2 Quasi-Newton-Methoden 43

4.3 Gauß-Hermite-Quadratur 45

5 DATENGRUNDLAGE DER EMPIRISCHEN ANALYSE DES ARBEITSSTATUS MÄNNLICHER ERWERBSPERSONEN 47

5.1 Definition der Variablen und deren erwarteter Einfluß auf den Arbeitsstatus 48

5.2 Besonderheiten des Datensatzes und daraus zu erwartende Schätzprobleme 51

6 ERGEBNISSE DER SCHÄTZUNG 53

6.1 Ein statisches Probitmodell für die einzelnen Wellen 53

6.2 Ein statisches Probitmodell unter der Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit 57

6.2.1 Ergebnisse der Schätzung eines statischen Probitmodells unter der Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit 60

6.2.2 Ergebnisse des Likelihood-Verhältnis-Tests auf Gleichheit der Parametervektoren für die Wellen B bis N 62

6.3 Ein statisches Probitmodell mit zufälligen, individuenspezifischen Effekten unter der Annahme äquikorrelierter Störgrößen 63

6.4 Ein statisches Probitmodell unter der Annahme autokorrelierter Störgrößen 66

6.5 Ein dynamisches Probitmodell unter der Annahme äquikorrelierter Störgrößen - Exogener Anfangszustand 71

6.6 Ermittlung marginaler Effekte für ausgewählte Probitmodelle 74

6.7 Güte der Anpassung in ausgewählten Probitmodellen 76

7 ZUSAMMENFASSUNG 80

8 LITERATUR 83

Tabellenverzeichnis

6.1.1 Statisches Probitmodell; Ergebnisse der Parameterschätzung für die einzelnen Wellen; Wellen B (1985) bis H (1991) 54

6.1.2 Statisches Probitmodell; Ergebnisse der Parameterschätzung für die einzelnen Wellen; Wellen I (1992) bis N (1997) 55

6.2.1 Veränderungen des Arbeitsstatus der Individuen beim Übergang von einer Welle zur nächsten; Wellen B (1985) bis N (1997) 58

6.2.2 Statisches Probitmodell unter der Annahme vollständiger Unabhängigkeit der individuellen Entscheidungen über die Zeit; Ergebnisse der Parameterschätzung; Wellen B (1985) bis N (1997)59

6.3.1 Statisches Probitmodell mit zufälligen, individuenspezifischen Effekten unter der Annahme äquikorrelierter Störgrößen; Ergebnisse der Parameterschätzung; Wellen B (1985) bis N (1997)64

6.4.1 Statisches Probitmodell unter der Annahme autokorrelierter Störgrößen; Ergebnisse der MECOSA-Schätzung der Kovarianzmatrix; Wellen E (1988) bis N (1997); Stufe 2 68

6.4.2 Statisches Probitmodell unter der Annahme autokorrelierter Störgrößen; Ergebnisse der Parameterschätzung für das Panel unter Verwendung des MECOSA-Algorithmus; Wellen E (1988) bis N (1997); Stufe 3 69

6.5.1 Dynamisches Probitmodell unter der Annahme äquikorrelierter Störgrößen - Exogener Anfangszustand; Ergebnisse der Parameterschätzung; Wellen B (1985) bis N (1997) 72

6.6.1 Ergebnisse der Schätzung der marginalen Effekte; Ausgewählte Panelmodelle; Wellen B (1985) bis N (1997) 75

6.7.1 Ergebnisse der Schätzung für die Güte der Anpassung; Ausgewählte Panelmodelle; Wellen B (1985) bis N (1997) 78

1 Einleitung

1.1 Problemstellung

Seit ungefähr 25 Jahren (vgl. das Vorwort in Ronning, 1991) werden in der Ökonometrie - und dies in neuerer Zeit in zunehmendem Maße - Individualdaten verfügbar und ausgewertet, nachdem bis dahin die Analyse von Aggregaten in der empirischen Wirtschaftsforschung im Vordergrund stand. So gelingt es z.B. statt einer Analyse des Aggregates ′Arbeitslosenquote′ nunmehr, Faktoren herauszuarbeiten, welche maßgebend sind dafür, ob eine bestimmte Person arbeitslos ist oder nicht.

In den vergangenen Jahren haben dabei Analysen auf der Basis von Panelerhebungen zunehmend an Bedeutung gewonnen. Man spricht von Paneldaten, wenn für eine bestimmte, gleichbleibende Gruppe von Untersuchungseinheiten für mehrere aufeinanderfolgende Zeitpunkte Beobachtungen für dasselbe Merkmal vorliegen.

Mit dem zunehmenden Grad an Komplexität des bei Panelerhebungen gewonnenen Datenmaterials hat jedoch die Entwicklung und Implementierung geeigneter ökonometrisch-statistischer Methoden in einigen Bereichen nicht Schritt gehalten. Dieses gilt u.a. für die Schätzung nichtlinearer Panelmodelle mit qualitativen abhängigen Variablen, und hier insbesondere für dynamische nichtlineare Panelmodelle. Auf diesem Gebiet besteht demzufolge ein erheblicher Forschungsbedarf. Deshalb ist ein besonderer Schwerpunkt der hier vorgelegten Arbeit in der quantitativen Analyse dynamischer nichtlinearer Panelmodelle zu sehen.

Der vergleichsweise hohe Komplexitätsgrad derartiger Modelle führt aber dazu, daß in der vorgelegten Arbeit nur solche Modelle für Paneldaten behandelt werden, in denen die endogene Variable einer binären (dichotomen) Zufallsvariablen entspricht (sog. binomiale Modelle). Als abhängige Variable kommen deshalb wirtschaftliche Größen, wie z.B. der Beschäftigungsstatus (beschäftigt/arbeitslos), in Betracht. Dieses Vorgehen führt zu Modellen, die z.B. eine Analyse individueller Arbeitslosigkeit ermöglichen.

Eine Fragestellung wie diese legt nahe, daß das zu betrachtende nichtlineare Panelmodell als erklärende Variable u.a. auch verzögert endogene Variable enthalten sollte, d.h., ein dynamisches nichtlineares Panelmodell zu unterstellen. Eine derartige Dynamisierung im genannten Beispiel erscheint sinnvoll, da die ′heute′ bestehende Arbeitslosigkeit einer betrachteten Person offenbar auch einen Einfluß auf deren Wahrscheinlichkeit, ′morgen′ arbeitslos zu sein, besitzt.

Die Formulierung der nichtlinearen Panelmodelle erfolgt auch in der hier vorgelegten Arbeit auf der theoretischen Grundlage eines Schwellenwertmodells. Dies bedeutet: Überschreitet eine latente (nicht beobachtbare) Zufallsvariable einen vorgegebenen Schwellenwert (i.a. den Wert null), so nimmt die zugehörige (beobachtbare) endogene Variable den Wert eins an; bleibt die latente Zufallsvariable dagegen unterhalb des Schwellenwertes, so ergibt sich für die endogene Variable der Wert null. Im oben genannten Beispiel entspricht die latente Zufallsvariable der Disposition einer Person, arbeitslos zu sein bzw. zu werden und kann in diesem Zusammenhang als Differenz von Anspruchslohn und Marktlohn der betrachteten Person interpretiert werden (vgl. Heckman, 1981a, S. 120); als Schwellenwert bietet sich demzufolge der Wert null an. Die endogene Variable wird auf den Wert eins gesetzt, falls die betrachtete Person arbeitslos ist; andernfalls ist die endogene Variable gleich null.

Hinsichtlich der Spezifikation derartiger Modelle ergibt sich folgendes: Im einfachsten Fall wird man nicht autokorrelierte Störgrößen unterstellen; da diese Annahme jedoch häufig realitätsfern ist (sie geht von vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit aus), werden in einem weiteren Schritt Panelmodelle analysiert, die von der Annahme einer Äquikorrelation der Störgrößen für ein bestimmtes Individuum über die Zeit ausgehen; diese Struktur der Störgrößen ergibt sich bei Einbeziehung stochastischer, individuenspezifischer Effekte in das Schwellenwertmodell. Schließlich wird versucht, auch komplexere Autokorrelationsprozesse der Störterme, beispielsweise einen Markov-Prozeß erster Ordnung, in dem Modell zu berücksichtigen.

1.2 Zielsetzung und Vorgehensweise

Das Ziel dieser Arbeit ist demzufolge in einer Anwendung und spezifischen Weiterentwicklung bestehender ökonometrisch-statistischer Methoden bezogen auf eine konkrete Fragestellung aus dem Bereich dynamischer Panelmodelle mit binären abhängigen Variablen zu sehen.

Für eine empirische Implementierung der hier behandelten methodischen Ansätze werden Daten des Deutschen Sozio-ökonomischen Panels (SOEP) herangezogen und zwar Angaben aus den bisher vorliegenden Wellen A bis N dieses Panels (1984 bis 1997). Analysiert werden Ursachen der individuellen Arbeitslosigkeit auf dem deutschen Arbeitsmarkt und zwar von Männern (Arbeitern und Angestellten) mit einem Eintrittsalter von 25 bis 51 Jahren (im Basisjahr 1984).

Für eine quantitative Analyse nichtlinearer Panelmodelle ist es wünschenswert, dynamische Panelmodelle mit beliebiger Korrelationsstruktur der Störgrößen zu betrachten. Dieses erfordert, sowohl verzögert endogene Variable als erklärende Variable in das Panelmodell aufzunehmen, als auch eine beliebige Korrelationsstruktur für die latenten Zufallsvariablen über die

Zeit zuzulassen. Es stellt sich dann allerdings die Frage, inwieweit eine solche wünschenswerte Modellspezifikation auch schätzbar ist.

Zunächst besteht also in jedem Fall nicht nur eine Unsicherheit bzgl. der Variablenwahl sondern auch eine Unsicherheit bzgl. der Spezifikation der Panelmodelle; zudem erfordern unterschiedliche Modellansätze eventuell auch unterschiedliche Verfahren für die Parameterschätzung. Der Vollständigkeit halber sei aber ergänzt, daß alle in dieser Arbeit verwendeten Parameterschätzverfahren auf der unmodifizierten Maximum Likelihood-Methode (ML-Methode) bzw. auf Modifikationen dieser Methode beruhen.

Darüber hinaus wird bei der empirischen Analyse in allen hier betrachteten (Schwellenwert-) Modellen von (standard-) normalverteilten Störgrößen ausgegangen, d.h., es werden gemäß der getroffenen Annahmen, in allen Fällen Probitmodelle unterstellt.

Ausgangspunkt der Panelanalyse sind folglich Schätzungen eines statischen binomialen Antwortmodells auf der Basis von Querschnittsdaten für die einzelnen Wellen B bis N (1985 bis 1997)[1].

In einem zweiten Schritt wird ein Panelmodell in Form eines statischen Probitmodells über alle dreizehn Wellen des Panels geschätzt. Dieses Modell geht von vollkommener Homogenität der individuellen Entscheidungen über die Zeit aus. Mittels der Werte der logarithmierten Likelihoodfunktionen (an der Stelle des zugehörigen ML-Schätzers) für die einzelnen Wellen und dem Wert der logarithmierten Likelihoodfunktion für das Probitmodell über alle dreizehn Wellen des Panels wird dann ein Test auf Gleichheit der dreizehn Parametervektoren (der einzelnen Wellen) formuliert.

Die Annahme vollkommener Homogenität der individuellen Entscheidungen über die Zeit wird anschließend modifiziert, indem zusätzlich stochastische, individuenspezifische Effekte explizit im Modell berücksichtigt werden. Diese Spezifikation der zufälligen Effekte unterstellt im einfachsten Fall eine Äquikorrelation der Störgrößen (das sog. Varianzkomponentenmodell). Auch diese Hypothese ist Gegenstand eines entsprechenden Tests.

In weiteren Schritten werden die obigen Modelle einerseits durch die Einbeziehung verzögert endogener Variablen zwecks Berücksichtigung möglicher Zustandsabhängigkeiten erweitert und andererseits durch die Spezifikation komplexerer Störgrößenprozesse verallgemeinert. In diesen dynamischen Probitmodellen wird dann auf Zustandsabhängigkeit bzw. auf Autokorrelation der Störgrößen getestet.

---

[1] Bei der Implementierung der Panelmodelle wird die Welle A nicht berücksichtigt, da eine wichtige Modellvariable erst ab der Welle B verfügbar ist.

Aus dieser Vorgehensweise ergibt sich die folgende Gliederung der vorliegenden Arbeit:

Gegenstand der Ausführungen des zweiten Kapitels ist die Spezifikation statischer nichtlinearer Panelmodelle. Im Anschluß an einen Hinweis auf die Vorteile der Verwendung von Paneldaten (Abschnitt 2.1) stehen die Herleitung eines nichtlinearen Panelmodells mittels eines Schwellenwertmodells und die Spezifikation der Störgrößen in derartigen Modellen (feste Effekte versus zufällige Effekte) im Vordergrund (Abschnitt 2.2).

Zur Erläuterung wird im Abschnitt 2.2.1 ein Logitmodell mit festen Effekten behandelt, wobei die Schätzung der Strukturparameter dieses Modells mittels einer bedingten Likelihoodfunktion erläutert wird. Der Abschnitt 2.2.2 befaßt sich mit der Spezifikation zufälliger Effekte in nichtlinearen Panelmodellen. Dabei werden die diesbezüglichen Möglichkeiten anhand eines statischen Probitmodells mit äquikorrelierten Störgrößen (Abschnitt 2.2.2.1) und des Chamberlain-Ansatzes (Abschnitt 2.2.2.2) erläutert.

Verschiedene Aspekte der ML-Schätzung in binomialen Antwortmodellen (insbesondere Probitmodellen) werden in Abschnitt 2.3 behandelt. Der Abschnitt 2.3.1 skizziert die Newton-Raphson-Methode. Methoden zur Schätzung der asymptotischen Kovarianzmatrix des ML-Schätzers werden im Abschnitt 2.3.2 erläutert. Der Schwerpunkt liegt dabei auf der von Berndt, Hall, Hall und Hausman entwickelten (BHHH)-Methode. Abschnitt 2.3.3 verweist auf den im empirischen Teil der Arbeit verwendeten Likelihood-Quotienten-Test.

Gegenstand der Ausführungen des dritten Kapitels sind dynamische nichtlineare Panelmodelle. In den Vorbemerkungen zu diesem Kapitel (Abschnitt 3.1) werden die Begriffe ′Zustandsabhängigkeit′ und ′unbeobachtete Heterogenität′ (′unobserved heterogeneity′; siehe Hamerle und Ronning, 1994, S. 412) erläutert und gegeneinander abgegrenzt.

Ein Modell zur Spezifikation dynamischer nichtlinearer Panelmodelle in Anlehnung an Heckman (Abschnitt 3.2) stellt die Grundlage der in den dann folgenden drei Abschnitten betrachteten Modelle dar: eines statischen Probitmodells mit autokorrelierten Störgrößen[1] (Abschnitt 3.2.1), eines dynamischen Probitmodells mit äquikorrelierten Störgrößen (Abschnitt 3.2.2) und eines dynamischen Probitmodells mit autokorrelierten Störgrößen (Abschnitt 3.2.3).

Die Aussagen des Abschnitts 2.3 bezüglich der Implementierung statischer und dynamischer Panelmodelle werden in Kapitel 4 erweitert bzw. vertieft.

---

[1] Das statische Probitmodell mit autokorrelierten Störgrößen wird in der vorliegenden Arbeit, aufgrund der getroffenen Annahmen bezüglich der Dynamik des Störgrößenprozesses (eines Markov-Prozesses erster Ordnung), zu den dynamischen Probitmodellen gezählt, obwohl keine verzögert endogenen Variablen als erklärende Variable in das Modell eingehen.

Dies ist notwendig, da eine Autokorrelation der Störgrößen im Probitmodell impliziert, daß die zugehörige logarithmierte Likelihoodfunktion bei T Beobachtungszeitpunkten T-dimensionale Normalverteilungsintegrale enthält, die bereits für kleinere Werte von T zumindest schwierig auszuwerten sind. Im Abschnitt 4.1 wird dazu der MECOSA-Ansatz betrachtet, der eine mögliche Lösung dieses Problems liefert.

Der Abschnitt 4.2 greift zwei spezielle iterative Verfahren der nichtlinearen Optimierung heraus, den Davidon-Fletcher-Powell (DFP)- und den Broyden-Fletcher-Goldfarb-Shanno (BFGS)-Algorithmus. Beide Algorithmen liefern Weiterentwicklungen der Newton-Raphson-Methode, und benötigen im Gegensatz zu dieser für die Maximierung der jeweiligen logarithmierten Likelihoodfunktion ausschließlich deren partielle erste Ableitungen nach dem zugehörigen Parametervektor.

Schließlich wird auf folgendes Problem eingegangen: Die logarithmierte Likelihoodfunktion im statischen bzw. im dynamischen Probitmodell mit äquikorrelierten Störgrößen enthält Integrale von Produkten über T Standardnormalverteilungen (vgl. Abschnitte 2.2.2.1 und u.a. 3.2.2). Zur numerisch effizienten Berechnung derartiger Integrale eignet sich u.a. die Gauß-Hermite-Quadratur, die in Abschnitt 4.3 behandelt wird.

Kapitel 5 liefert die notwendigen Angaben über die Datengrundlage der anschließenden empirischen Analyse der Ursachen individueller Arbeitslosigkeit männlicher Erwerbspersonen in der Bundesrepublik Deutschland. Diese Analyse erfolgt auf der Basis von Daten des Deutschen Sozio-ökonomischen Panels. Das schließlich verwendete Datenpanel enthält Angaben für 527 Personen über 13 Wellen. Dazu werden in Abschnitt 5.1 die in den verschiedenen Modellen berücksichtigten Variablen definiert und deren erwarteter Einfluß auf den Arbeitsstatus herausgearbeitet; im Abschnitt 5.2 werden die Besonderheiten des Datensatzes und bei der empirischen Analyse der nichtlinearen Panelmodelle zu erwartende Schätzprobleme angesprochen.

In Kapitel 6 wird dieses Datenmaterial, unter Verwendung der verschiedenen statischen und dynamischen Probitmodelle analysiert. Die empirische Analyse der Ursachen individueller Arbeitslosigkeit beginnt in Abschnitt 6.1 mit der Schätzung der Strukturparameter in Probitmodellen für die einzelnen Wellen (Tabellen 6.1.1 und 6.1.2).

Gegenstand des Abschnitts 6.2 ist die Analyse des Panels über 13 Wellen (Wellen B bis N) mittels eines statischen Probitmodells, eines Modells, das von vollkommener Homogenität der individuellen Entscheidungen über die Zeit ausgeht (Tabelle 6.2.2). Dieser Schätzung vorgeschaltet, ist ein Überblick über das Ausmaß der Veränderungen des Arbeitsstatus der Individuen beim Übergang von einer Welle zur nächsten (Tabelle 6.2.1). Die Ausführungen dieses Abschnitts werden ergänzt durch einen Likelihood-Quotienten-Test auf Parameterkonstanz über die Zeit.

Ein statisches Panelmodell (Probitmodell) mit zufälligen Effekten (mit äquikorrelierten Störgrößen) liegt den Ausführungen des Abschnitts 6.3 zugrunde. Mit diesem Modell wird unter anderem getestet, ob die Annahme vollkommener Homogenität der individuellen Entscheidungen über die Zeit zutrifft, d.h., ob von über die Zeit unabhängigen Störgrößen ausgegangen werden kann oder nicht (Tabelle 6.3.1).

Im Abschnitt 6.4 wird ein alternativer Algorithmus, der sogenannte MECOSA-Ansatz zur Schätzung eines statischen Panelmodells (Probitmodells) mit autokorrelierten Störgrößen verwendet (Tabellen 6.4.1 und 6.4.2). Wesentliche Gesichtspunkte, die die Anwendung dieses Algorithmus nahelegen, sind Hinweise auf eine Autokorrelation der Störgrößen; diese Hypothese wird mittels des MECOSA-Ansatzes getestet.

Den Ausführungen des Abschnitts 6.5 liegt ein dynamisches Panelmodell (Probitmodell) mit äquikorrelierten Störgrößen und exogenem Anfagszustand zugrunde (Tabelle 6.5.1). Bei der empirischen Analyse dieses dynamischen Probitmodells wird auf Zustandsabhängigkeit getestet; in diesem Modell geschieht dies durch die Beantwortung der Frage, ob der um eine Periode verzögerte Arbeitsstatus einen signifikanten Einfluß auf die Wahrscheinlichkeit, beschäftigt zu sein, ausübt oder nicht.

Der Abschnitt 6.6 befaßt sich mit der Schätzung der marginalen Effekte in ausgewählten statischen und dynamischen Panelmodellen (Tabelle 6.6.1).

Im Abschnitt 6.7 wird die Güte der Anpassung in ausgewählten Probitmodellen betrachtet (Tabelle 6.7.1).

Kapitel 7 liefert eine Zusammenfassung der Ergebnisse der vorliegenden Arbeit.

2 Statische nichtlineare Panelmodelle

2.1 Einführende Bemerkungen zur Verwendung von Paneldaten

Mittels Paneldaten ist es im allgemeinen möglich, komplexere Verhaltensmodelle zu konstruieren, zu schätzen und zu testen, als dies mit Querschnittsdaten, einer Folge von Querschnitten oder einer Zeitreihe möglich wäre.
Neben einer erweiterten Datenbasis, die zu einer größeren Effizienz der ökonometrischen Schätzungen führen kann, bieten Paneldaten dem Ökonometriker hauptsächlich folgende Vorteile (vgl. Hsiao, 1985, S. 3) :

- die Identifikation ökonomischer Modelle und die Diskriminierung zwischen verschiedenen ökonomischen Hypothesen,
- die Eliminierung bzw. Reduktion der Verzerrung der Parameterschätzwerte, sowie
- die Möglichkeit der Reduktion des Multikollinearitätsproblems.

Auf diese drei Punkte soll nun näher eingegangen werden.

Paneldaten besitzen - gegenüber einer zeitlichen Folge von Querschnittsdaten - den (Informations-)Vorteil, daß dynamische Entwicklungen beobachtet werden können. Dazu folgendes Beispiel:

- Angenommen, die Arbeitslosenquote in Niedersachsen betrage in jedem Monat eines bestimmten Jahres $8,\overline{3}$ Prozent, so kann dieses Beobachtungsergebnis in einem Fall dadurch verursacht werden, daß $8,\overline{3}$ Prozent der Personen über alle zwölf Monate des Jahres arbeitslos sind, zum anderen aber auch dadurch, daß jedes Individuum der beobachteten Grundgesamtheit jeweils genau einen Monat keiner Beschäftigung nachgeht. Eine Unterscheidung dieser beiden Extreme, sowie eine Abgrenzung gegenüber einer Vielzahl weiterer Varianten ist auf der Grundlage einer Folge von Querschnittsdaten allein nicht möglich. Dagegen liefern Paneldaten z.B. die Information, wieviele der zum Zeitpunkt t vorhandenen Arbeitslosen im Zeitpunkt $t+1$ eine Beschäftigung fanden und eröffnen somit gegebenenfalls die Möglichkeit, Ursachen und Wirkungen einander zuzuordnen.
- Ein weiteres, für die vorliegende Fragestellung relevantes Beispiel, welches die Vorteile der Verwendung von Paneldaten bei der Diskriminierung konkurrierender ökonomischer Theorien zeigt, ergibt sich aus einer Theorie zur Entstehung von Arbeitslosigkeit. Gilt es, die Frage zu klären, ob die Linderung der gegenwärtigen Arbeitslosigkeit einen Effekt auf die Höhe der zukünftigen Arbeitslosigkeit besitzt, dann hängt deren Beantwortung eng

damit zusammen, ob eine bestehende Arbeitslosigkeit eines Individuums dessen Wahrscheinlichkeit erhöht, auch in den Folgeperioden arbeitslos zu sein oder nicht. Die Verwendung von Paneldaten eröffnet die Möglichkeit, auf derartige Zustandsabhängigkeiten zu testen und diese gegebenenfalls im Rahmen der Modellspezifikation zu berücksichtigen (vgl. Flaig, Licht und Steiner, 1992).

Die Vorteile der Verwendung von Paneldaten im Hinblick auf die Eliminierung einer Verzerrung der Parameterschätzwerte kann man anhand eines einfachen linearen Regressionsmodells zeigen. Ein Panelmodell für N Individuen und T Wellen sei gegeben durch

$$y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{z}_{it}'\boldsymbol{\gamma} + u_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T.$$

Dabei bezeichne y_{it} die endogene Variable für das Individuum i zum Zeitpunkt t, $\boldsymbol{x}_{it}$ und $\boldsymbol{z}_{it}$ seien K-dimensionale Spaltenvektoren exogener Variablen, $\boldsymbol{\beta}$ und $\boldsymbol{\gamma}$ die zugehörigen K-dimensionalen Parametervektoren und u_{it} die Störgröße für das Individuum i zum Zeitpunkt t.

Sind die exogenen Variablen $\boldsymbol{x}_{it}$ und $\boldsymbol{z}_{it}$ nichtstochastisch und die Störgrößen u_{it} stochastisch unabhängig und identisch verteilt über i und t, so liefert die KQ-Regression von y_{it} auf $\boldsymbol{x}_{it}$ und $\boldsymbol{z}_{it}$ unverzerrte und konsistente Schätzer für die Parametervektoren $\boldsymbol{\beta}$ und $\boldsymbol{\gamma}$.

Handelt es sich bei den $\boldsymbol{z}_{it}$ hingegen um (unbeobachtete) stochastische Variablen, die mit den exogenen Variablen $\boldsymbol{x}_{it}$ korreliert sind, dann liefert z.B. die KQ-Schätzung für den Parametervektor $\boldsymbol{\beta}$ verzerrte Schätzwerte.

Kann jedoch bei Vorliegen von Paneldaten unterstellt werden, daß diese $\boldsymbol{z}_{it}$ additive, zeitunabhängige individuenspezifische Effekte darstellen, d.h., gilt $\boldsymbol{z}_{it} = \boldsymbol{z}_i$ für alle t (das Modell lautet dann: $y_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \boldsymbol{z}_i'\boldsymbol{\gamma} + u_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T$), so kann man auch in diesem Modell zu einer unverzerrten und konsistenten Schätzung des Parametervektors $\boldsymbol{\beta}$ gelangen (vgl. Hsiao, 1986, S. 31); denn z.B. durch Transformation der individuellen Beobachtungen in die Abweichungen von ihrem individuenspezifischen Mittelwert (über die Zeit), lassen sich die additiven, individuenspezifischen Effekte $\boldsymbol{z}_i$ eliminieren. Man gelangt somit zu folgendem Regressionsmodell:

$$y_{it} - \bar{y}_i = (\boldsymbol{x}_{it} - \bar{\boldsymbol{x}}_i)'\boldsymbol{\beta} + (u_{it} - \bar{u}_i), \quad i = 1,\ldots,N;\ t = 1,\ldots,T,$$

mit

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \quad \bar{\boldsymbol{x}}_i = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_{it} \quad \text{und} \quad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{it}$$

für das z.B. die KQ-Schätzung unverzerrte und konsistente Schätzwerte für den Parametervektor $\boldsymbol{\beta}$ liefert (vgl. Hsiao, 1986, S. 4).

Hinsichtlich der sich bei der empirischen Analyse von Zeitreihendaten häufig ergebenden Multikollinearitätsprobleme, welche eine hinreichend genaue Bestimmung der individuellen Einflüsse der erklärenden Variablen erschweren oder diese sogar unmöglich machen, liegt der Vorteil der Verwendung von Paneldaten insbesondere in einer Datenbasis größerer Variabilität gepaart mit dem Vorhandensein individuenspezifischer Merkmale. Beides führt im allgemeinen zu einer Reduktion des Multikollinearitätsproblems.

2.2 Grundlagen der Spezifikation statischer nichtlinearer Panelmodelle

Im folgenden geht es um die Spezifikation statischer nichtlinearer Panelmodelle, für N Individuen über T (aufeinanderfolgende) Wellen. Es werden wie bereits angeführt, nur binomiale Modelle für Paneldaten behandelt. Für die endogene Variable, eine binäre Zufallsvariable, gelte beispielsweise:

$$y_{it} = \begin{cases} 1 & \text{falls das Individuum } i \text{ zum Zeitpunkt } t \text{ arbeitslos ist} \\ 0 & \text{sonst.} \end{cases}$$

Um diesen Gedanken fortzusetzen, wird nun argumentiert, daß bestimmte Faktoren einen Einfluß auf die Wahrscheinlichkeit, arbeitslos zu sein bzw. zu werden, besitzen. Diese Faktoren seien in einem Vektor exogener Variablen $\boldsymbol{x_{it}}, i = 1,\ldots,N;\ t = 1,\ldots,T$ zusammengefaßt, mit anderen Worten, man interessiert sich vor allem dafür, welchen Einfluß eine Änderung bestimmter exogener Variablen auf die Wahrscheinlichkeit des Eintreffens eines vorgegebenen Ereignisses (Arbeitslosigkeit) besitzt.

Ein ′natürliches Gerüst′ (′natural framework′; Heckman, 1981a, S. 119) zur Formulierung entscheidungstheoretischer ökonometrischer Modelle schafft ein, dem binomialen Antwortmodell zugrundeliegendes, Schwellenwertmodell. Diesem Modell liegt folgender Gedanke zugrunde: Überschreitet eine latente (nicht beobachtbare) Zufallsvariable y_{it}^* einen Schwellenwert (i.a. den Wert null), so nimmt die beobachtbare endogene Variable y_{it} den Wert eins an; unterschreitet y_{it}^* den Schwellenwert, so nimmt y_{it} den Wert null an. Die latente stetige Zufallsvariable y_{it}^*, welche selbst nicht beobachtet werden kann, wird bei der Betrachtung von Paneldaten folgendermaßen definiert:

$$y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T.$$

Dabei bezeichnen y_{it}^* die latente Zufallsvariable für das i-te Individuum zum Zeitpunkt t, $\boldsymbol{x_{it}}$ einen entsprechenden K-elementigen Vektor exogener Variablen, $\boldsymbol{\beta}$ den K-elementigen Vektor der Strukturparameter und u_{it} die zugehörige Störgröße.

Die latente Zufallsvariable y_{it}^* entspricht, um bei dem o.a. Beispiel zu bleiben, der Disposition eines Individuums, arbeitslos zu sein bzw. zu werden und ist in diesem Zusammenhang z.B. als Differenz von Anspruchslohn und Marktlohn des Individuums i zum Zeitpunkt t interpretierbar (vgl. Heckman, 1981a, S. 120), wobei zumindest der Anspruchslohn des Individuums als stochastisch bzw. nicht beobachtbar einzustufen ist. Übersteigt der Anspruchslohn der betrachteten Person seinen Marktlohn (für eine bestimmte Tätigkeit), so wird die betrachtete Person diese Tätigkeit nicht annehmen und damit arbeitslos sein.

Auf der theoretischen Grundlage des Schwellenwertmodells stellt man nun folgenden formalen Zusammenhang zwischen der beobachtbaren binären Zufallsvariablen y_{it} und der latenten (nicht beobachtbaren) Zufallsvariablen y_{it}^* her:

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > \text{Schwellenwert (z.B. } y_{it}^* > 0) \\ 0 & \text{sonst.} \end{cases}$$

Im Hinblick auf das gewählte Beispiel des Arbeitsmarktes, kann der Schwellenwert, wie oben ausgeführt, als Differenz von Anspruchslohn und gebotenem Arbeitslohn interpretiert und somit gleich null gesetzt werden. Für die Wahrscheinlichkeit, daß y_{it} gleich eins ist, erhält man somit:

$$P(y_{it}=1) = P(y_{it}^*>0) = P(\boldsymbol{x_{it}'\beta} + u_{it} > 0) = P(u_{it} > -\boldsymbol{x_{it}'\beta}) .$$

Unterstellt man, daß die Störgröße u_{it} einer bestimmten, symmetrischen und standardisierten Verteilung mit der Verteilungsfunktion F genügt, so gilt:

$$P(y_{it}=1) = P(u_{it} < \boldsymbol{x_{it}'\beta}) = F(\boldsymbol{x_{it}'\beta}) .$$

Damit ist der Strukturparametervektor $\boldsymbol{\beta}$ identifizierbar.

Unterstellt man hingegen nur, daß die Störgröße u_{it} symmetrisch verteilt (z.B. normalverteilt) ist mit Erwartungswert null und beliebiger Varianz σ^2, so ergibt sich für die Wahrscheinlichkeit, daß die endogene Variable y_{it} den Wert eins annimmt, folgender Ausdruck:

$$P(y_{it}=1) = P(u_{it} < \boldsymbol{x_{it}'\beta}) = F\left(\frac{\boldsymbol{x_{it}'\beta}}{\sigma}\right) .$$

In diesem Fall ist nur der Term $\boldsymbol{\beta}/\sigma$ identifizierbar.

Für die Verteilungsfunktion F wird in den Panelmodellen entweder eine Standardnormalverteilung oder eine logistische Verteilung unterstellt, was zum Probit- bzw. Logitmodell führt.

Im einfachsten Fall unterstellt man, daß die u_{it} stochastisch unabhängig und identisch verteilt über alle i und t seien. Unter diesen Annahmen ergibt sich folgende Likelihoodfunktion für die Beobachtungen $y_{it}, i = 1,...,N; t = 1,...,T$ eines Panelmodells:

$$L = \prod_{i=1}^{N} \prod_{t=1}^{T} [F(\boldsymbol{x}_{it}'\boldsymbol{\beta})]^{y_{it}} [1 - F(\boldsymbol{x}_{it}'\boldsymbol{\beta})]^{1-y_{it}}$$

und die logarithmierte Likelihoodfunktion ist

$$ln\, L = \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it}\, ln\, F_{it} + (1 - y_{it})\, ln\, (1 - F_{it})),$$

mit $F_{it} := F(\boldsymbol{x}_{it}'\boldsymbol{\beta})$.

Paneldaten können aber durch komplexe Strukturen der Störterme charakterisiert sein. Deshalb gilt es bei der Analyse von Paneldaten, zusätzliche Annahmen über die Abhängigkeitsstruktur der u_{it} zu treffen. Unterstellt man nämlich, daß die u_{it} im obigen Schwellenwertmodell stochastisch unabhängig und identisch verteilt über i und t sind, so impliziert diese Annahme, daß die Koeffizienten für alle Individuen und über alle Zeitpunkte konstant sind. Diese, nicht sehr realitätsnahe, Annahme der Unabhängigkeit der Störgrößen für die jeweiligen Individuen über die Zeit bzw. vollkommener Homogenität der individuellen Entscheidungen über die Zeit führt jedoch zu inkonsistenten oder sinnlosen Schätzungen für die Strukturparameter, wenn diese Voraussetzungen nicht erfüllt sind.

Der Hauptvorteil der Verwendung von Paneldaten besteht aber gerade darin, daß einige der restriktiven Annahmen, wie diese für die Analyse von Querschnittsdaten erforderlich sind, bei der Verwendung von Paneldaten aufgehoben bzw. abgeschwächt werden können. So ist es z.B. möglich, nicht beobachtbare individuen- bzw. zeitspezifische Effekte, eine sogenannte ′unbeobachtete Heterogenität′ in Panelmodellen explizit zu modellieren; mittels Paneldaten läßt sich, wie Hamerle und Ronning dieses ausdrücken, die individuelle Heterogenität ′kontrollieren′ (Hamerle und Ronning, 1994, S. 412).

Für die Mehrzahl der zu betrachtenden Panelmodelle ist dies ein Mittel, um konsistente Schätzungen für den systematischen Teil des Modells, d.h., für den Strukturparametervektor $\boldsymbol{\beta}$, zu erhalten. Dieser Zusammenhang bedarf einer formalen Präzisierung.

Wird für das Schwellenwertmodell

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst,} \end{cases}$$

mit $y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_{it}$, $i = 1,\ldots,N;\ t = 1,\ldots,T$

und $y_{it}^*, \boldsymbol{x}_{it}, \boldsymbol{\beta}$ und u_{it} wie oben definiert (vgl. S. 9) unterstellt, daß die u_{it} stochastisch unabhängig und identisch verteilt über alle i und t sind, so spezifiziert man ein herkömmliches binomiales Antwortmodell mit $N \cdot T$ Beobachtungen, dessen Parameter man mit der ML-Methode, falls die entsprechenden Voraussetzungen erfüllt sind, konsistent und effizient schätzen kann (vgl. Abschnitt 2.3).

Diese Annahme vollkommener Unabhängigkeit der Entscheidungen aller einbezogenen Individuen über die Zeit kann sowohl theoretisch als auch empirisch nur selten gerechtfertigt werden. So können unbeobachtete bzw. unbeobachtbare individuenspezifische Effekte zu Autokorrelationen der Störgrößen und somit zu Verzerrungen der Parameterschätzwerte führen, soweit solche Effekte nicht im Rahmen der Schätzung berücksichtigt werden. Derartige Effekte lassen sich aber in Panelmodellen explizit modellieren. Hierzu kann die Störgröße u_{it} in drei Komponenten zerlegt werden:

$$u_{it} = \alpha_i + \lambda_t + \varepsilon_{it},$$

mit α_i, einem individuenspezifischen Effekt, λ_t, einem zeitspezifischen Effekt und ε_{it}, einer stochastisch unabhängigen und über alle Individuen und Wellen identisch verteilten Störgröße.

Hinsichtlich der α_i und λ_t kann man nun folgende Strategien verfolgen: Faßt man α_i und λ_t als unbekannte Parameter auf, welche zusammen mit dem Vektor der Strukturparameter $\boldsymbol{\beta}$ zu schätzen sind, so hat man es mit einem Ansatz mit festen Effekten (′Fixed-effects-Ansatz′) zu tun; modelliert man hingegen diese Effekte als stochastisch, so kommt man zu einem Ansatz mit zufälligen Effekten (′Random-effects-Ansatz′). In diesem Fall sind eine Wahrscheinlichkeitsverteilung der als stochastisch unterstellten Effekte α_i zu spezifizieren, und die unbekannten Parameter dieser Verteilung (im einfachsten Fall also der Mittelwert und die Varianz) zu schätzen.

Da aber im Normalfall die Anzahl der Querschnittseinheiten N groß und die Anzahl der Panelwellen T eher klein ist, modelliert man die zeitspezifischen Effekte λ_t, soweit erforderlich, grundsätzlich als feste Effekte, indem man sie in den Strukturparametervektor $\boldsymbol{\beta}$ aufnimmt und den Vektor der Regressoren in geeigneter Weise um Dummy-Variablen erweitert (vgl. Hamerle und Ronning, 1994, S. 418).

Wann, aber, wird man die individuenspezifischen Effekte als fest und wann als zufällig modellieren? Um diese Frage zu klären, sollen die beiden Formulierungen zunächst verein-

heitlicht werden. Man unterstellt dabei von Anfang an, daß die individuenspezifischen Effekte stochastisch sind (vgl. Hsiao, 1985, S. 11).

Bei einem Modell mit festen Effekten kommt der Ökonometriker zu Schlußfolgerungen aufgrund einer bedingten Betrachtung der N individuellen Effekte, d.h., die Likelihoodfunktion ist eine bzgl. der α_i bedingte Likelihood. Dieses Vorgehen ist sinnvoll, wenn man sich für die individuellen Effekte einzelner (weniger) Merkmalsträger interessiert und die Unterschiede zwischen diesen explizit schätzen will.

Unterstellt man, andererseits, die individuenspezifischen Effekte α_i als zufällig, ist also nicht an Stichprobenunterschieden sondern an Schlußfolgerungen über die Charakteristika der Grundgesamtheit der Individuen interessiert, dann ist die entsprechende Likelihoodfunktion eine, bezüglich der individuenspezifischen Effekte, unbedingte bzw. marginale Likelihood.[1]

Nachfolgend wird zunächst auf statische nichtlineare Panelmodelle mit festen bzw. mit zufälligen Effekten eingegangen.

2.2.1 Statische nichtlineare Panelmodelle mit festen Effekten

Als Basis der statischen nichtlinearen Panelmodelle mit festen Effekten kommt infrage ein Schwellenwertmodell mit folgender Spezifikation für die latente Zufallsvariable y_{it}^*:

$$y_{it}^* = \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \varepsilon_{it}\,, \quad i = 1,\ldots,N;\, t = 1,\ldots,T\,,$$

mit $y_{it}^*, \boldsymbol{x}_{it}$ und $\boldsymbol{\beta}$ wie oben (vgl. S. 9) definiert, mit dem individuenspezifischen Effekt α_i und mit ε_{it} stochastisch unabhängig und über alle *i* und *t* identisch verteilt.

Probleme bereitet in derartigen nichtlinearen Panelmodellen die Einbeziehung der zu schätzenden individuenspezifischen Parameter α_i deshalb, weil die Anzahl der damit - zusätzlich zum Strukturparametervektor $\boldsymbol{\beta}$ zu schätzenden - Parameter mit der Anzahl der Individuen steigt und diese Parameter nicht konsistent geschätzt werden können, da die Anzahl der Panelwellen (T) im allgemeinen eher klein ist. Hinzu kommt, daß sich eine eventuell vorliegende Inkonsistenz der Schätzung dieser individuenspezifischen Parameter (α_i) hier - im Gegensatz zu den linearen Panelmodellen - auf die Schätzung des Vektors der Strukturparameter $\boldsymbol{\beta}$ überträgt (denn die ML-Schätzungen der $\boldsymbol{\beta}$ und α_i erfolgen nicht unabhängig voneinander (vgl. Hsiao, 1986, S. 159ff.)).

---

[1] Diese Betrachtungsweise ist dann sinnvoll, wenn die Anzahl der Querschnittseinheiten groß ist und man die Erhebungsdaten als eine Zufallsstichprobe aus einer umfangreichen Grundgesamtheit betrachten kann.

Die Übertragung der Inkonsistenz der Schätzwerte für α_i bei festem T auf die Schätzwerte für $\boldsymbol{\beta}$ zeigt Andersen (1973) anhand eines, allerdings sehr restriktiven, Beispiels (für $T = 2$ und einen Regressor, wobei der Regressor in der ersten Periode gleich null und in der zweiten Periode gleich eins ist). Andersen gelangt für dieses Beispiel zu dem Ergebnis, daß der ML-Schätzer für N gegen unendlich gegen $2\boldsymbol{\beta}$ konvergiert und somit nicht konsistent ist (zitiert in: Hsiao, 1986, S. 159ff.).

Für $T > 2$ konnte Andersen das Ausmaß der Inkonsistenz nicht explizit nachweisen; Simulationsstudien legen aber nahe, daß $\hat{\boldsymbol{\beta}}$ gegen $\frac{T}{T-1}\boldsymbol{\beta}$ konvergiert (vgl. Hamerle und Ronning, 1994, S. 414).

Eine Kovarianztransformation, wie im linearen Modell möglich, ist in nichtlinearen Panelmodellen jedoch nicht durchführbar. Andersen (1973) schlägt deshalb für diesen Fall eine Schätzung des Strukturparametervektors $\boldsymbol{\beta}$ auf der Basis einer bedingten Likelihoodfunktion vor, welche allerdings nur im Logitmodell zu sinnvollen Schätzern führt, da man im Probitmodell keine suffiziente Statistik für die individuenspezifischen Effekte α_i findet (zitiert in: Hsiao, 1986, S. 163).[1] Obwohl Logitmodelle hier nicht mehr verwendet werden, soll dieser Zusammenhang um der Anschauung des Problems willen, kurz dargelegt werden.

Diese von Andersen vorgeschlagene Schätzung des Strukturparametervektors $\boldsymbol{\beta}$ auf der Basis einer bedingten Likelihood fußt auf folgenden Überlegungen (vgl. Hsiao, 1986, S. 161f.):

Für ein Logitmodell mit festen Effeken auf der Basis des o.a. Schwellenwertmodells ergibt sich für die Wahrscheinlichkeit, daß die endogene Variable y_{it} den Wert eins annimmt folgender Zusammenhang:

$$P(y_{it} = 1) = F(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta}),$$

oder:

$$P(y_{it} = 1) = \frac{exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}{1 + exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})} = \frac{e_{it}}{1 + e_{it}}, \quad i = 1,\ldots,N;\, t = 1,\ldots,T,$$

mit $e_{it} = exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})$.

Die gemeinsame Wahrscheinlichkeit der y_{it} über alle $t = 1,\ldots,T$ ist somit:

---

[1] Eine Statistik ist suffizient für α_i, falls sie die gesamte bezüglich α_i relevante Information der Stichprobe enthält; damit ist die bedingte Verteilung des Vektors der endogenen Variablen $\boldsymbol{y_i}$, mit $\boldsymbol{y_i} = (y_{i1},\ldots,y_{iT})'$, bei gegebenem Wert der suffizienten Statistik für α_i, unabhängig von α_i (vgl. Rasch, 1976, S. 208).

$$P(\boldsymbol{y_i}) = \prod_{t=1}^{T} \left(P(y_{it}=1)\right)^{y_{it}} \left(P(y_{it}=0)\right)^{1-y_{it}} = exp\left(\sum_{t=1}^{T}\left(y_{it}\, ln\, P(y_{it}=1) + (1-y_{it})\, ln\, P(y_{it}=0)\right)\right),$$

mit $\boldsymbol{y_i} = (y_{i1},\ldots,y_{iT})'$.

Andererseits ist

$$P(y_{it}=0) = \frac{1}{1+e_{it}} \text{ und damit:}$$

$$P(\boldsymbol{y_i}) = exp\left(\sum_{t=1}^{T}\left(y_{it}(ln\, e_{it} - ln(1+e_{it})) + (1-y_{it})\cdot -ln(1+e_{it})\right)\right) = exp\left(\sum_{t=1}^{T}(y_{it}\, ln\, e_{it} - ln(1+e_{it}))\right)$$

und wegen $ln\, e_{it} = \alpha_i + \boldsymbol{x'_{it}\beta}$ erhält man letztlich:

$$P(\boldsymbol{y_i}) = exp\left(\alpha_i \sum_{t=1}^{T} y_{it} + \sum_{t=1}^{T} \boldsymbol{x'_{it}}\beta \cdot y_{it} - \sum_{t=1}^{T} ln(1+exp(\alpha_i + \boldsymbol{x'_{it}}\beta))\right).$$

Die Theorie der Verteilungen einer Exponentialfamilie besagt nun, daß $\sum_{t=1}^{T} y_{it}$ eine suffiziente Statistik für den individuenspezifischen Effekt α_i ist (vgl. Hamerle und Ronning, 1994, S. 414f.). Damit läßt sich, gegeben $\sum_{t=1}^{T} y_{it}$, die bedingte Wahrscheinlichkeit für $\boldsymbol{y_i}$ ermitteln, die unabhängig von α_i ist.
Die bedingte Verteilung für $\boldsymbol{y_i}$, hängt folglich nicht mehr von den individuenspezifischen Parametern α_i ab und kann damit verwendet werden, um den Strukturparametervektor $\boldsymbol{\beta}$ konsistent zu schätzen.

Für die bedingte Likelihoodfunktion zeigt sich nun, daß Individuen, die in allen Wellen den gleichen Zustand aufweisen, nicht zur bedingten Likelihoodfunktion beitragen, da deren zugehörige bedingte Wahrscheinlichkeit den Wert eins annimmt. So ist beispielsweise: $P(\boldsymbol{y'_i} = (1,\ldots,1) | \sum_t y_{it} = T) = 1$.

Für $T=2$ mit dem einzig relevanten Fall $\sum_{t=1}^{2} y_{it} = 1$, (d.h., die Individuen weisen einen von der ersten zur zweiten Welle veränderten Zustand auf) erhält man beispielsweise für die Wahrscheinlichkeit, daß ein Individuum vom Zustand null in den Zustand eins wechselt, folgenden Ausdruck (vgl. Hamerle und Ronning, 1994, S. 415):

$$P\left(\boldsymbol{y'_i} = (0,1) \,\middle|\, y_{i1} + y_{i2} = 1\right) = \frac{exp\left((\boldsymbol{x'_{i2}} - \boldsymbol{x'_{i1}})\beta\right)}{1 + exp\left((\boldsymbol{x'_{i2}} - \boldsymbol{x'_{i1}})\beta\right)}.$$

Dies zeigt, daß die Veränderungen der Merkmalsausprägungen der exogenen Variablen von $t = 1$ nach $t = 2$ den Zustandswechsel der Individuen von der ersten zur zweiten Welle erklären, und die bedingte Wahrscheinlichkeit nicht mehr von den individuenspezifischen Effekten (α_i) abhängt.

Damit gelangt man für den Fall zweier Wellen zu folgender bedingten logarithmierten Likelihoodfunktion $ln\ L_B$:

$$ln\, L_B = \sum_{i \in I} \left(h_i (\boldsymbol{x}'_{i2} - \boldsymbol{x}'_{i1}) \boldsymbol{\beta} - ln \left(1 + exp\left((\boldsymbol{x}'_{i2} - \boldsymbol{x}'_{i1}) \boldsymbol{\beta} \right) \right) \right),$$

mit $I = \left\{ i : \sum_{t=1}^{2} y_{it} = 1 \right\}$ und $h_i = \begin{cases} 1 & \text{falls } \boldsymbol{y_i} = (0, 1) \\ 0 & \text{falls } \boldsymbol{y_i} = (1, 0). \end{cases}$

Diese bedingte logarithmierte Likelihoodfunktion $ln\ L_B$, mit der binären Zufallsvariablen h_i und dem Vektor der exogenen Variablen $\boldsymbol{x}'_{i2} - \boldsymbol{x}'_{i1}$, entspricht der ′üblichen′ logarithmierten Likelihoodfunktion des Logitmodells und stellt deshalb keine besonderen Ansprüche hinsichtlich des Schätzalgorithmus.

2.2.2 Statische nichtlineare Panelmodelle mit zufälligen Effekten

Gegenstand dieses Abschnitts sind statische nichtlineare Panelmodelle mit zufälligen Effekten. Dabei wird zunächst unterstellt, daß die individuenspezifischen Parameter α_i stochastisch und unabhängig vom Vektor der exogenen Variablen $\boldsymbol{x_i}$ verteilt sind und einer univariaten Verteilung (in der Regel einer Normalverteilung) genügen (Abschnitt 2.2.2.1). In einem weiteren Schritt wird diese Annahme fallengelassen, indem eine Korrelation der α_i und der $\boldsymbol{x_i}$ nicht ausgeschlossen wird (Abschnitt 2.2.2.2).

Die multivariate logistische Verteilung ist für derartige Betrachtungen ungeeignet, da sie in der Regel nur die Modellierung einer Äquikorrelation erlaubt (vgl. Hamerle und Ronning, 1994, S. 418). Es wird deshalb nachfolgend ausschließlich eine multivariate Normalverteilung unterstellt, damit können auch komplexere Störgrößenprozesse modelliert werden. Die folgenden Ausführungen beschränken sich deshalb auf die Betrachtung von Probitmodellen.

2.2.2.1 Probitmodelle mit äquikorrelierten Störgrößen

Für die beobachtete Zufallsvariable y_{it} gelte das o.a. Schwellenwertmodell (vgl. S. 11) und für die latente Zufallsvariable y^*_{it}:

$$y^*_{it} = \boldsymbol{x}'_{it} \boldsymbol{\beta} + u_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T,$$

mit $y^*_{it}, \boldsymbol{x_{it}}, \boldsymbol{\beta}$ und u_{it} wie oben (vgl. S. 9) definiert.

Zusammengefaßt über alle T Zeitpunkte ergibt sich hieraus:

$$\boldsymbol{y_i^*} = \boldsymbol{X_i}\,\boldsymbol{\beta} + \boldsymbol{u_i}, \quad i = 1,\ldots,N,$$

mit $\boldsymbol{y_i^*} = (y_{i1}^*,\ldots,y_{iT}^*)'$ dem T-dimensionalen Vektor der latenten Zufallsvariablen, $\boldsymbol{X_i} = (\boldsymbol{x_{i1}},\ldots,\boldsymbol{x_{iT}})'$ der $(T \times K)$-Matrix der exogenen Variablen, $\boldsymbol{\beta}$ dem K-dimensionalen Vektor der Strukturparameter und $\boldsymbol{u_i} = (u_{i1},\ldots,u_{iT})'$, dem T-dimensionalen Vektor der Störgrößen.

Für die Störgröße u_{it} gelte das folgende Varianzkomponentenmodell:

$$u_{it} = \alpha_i + \varepsilon_{it}, \quad \text{für alle } i = 1,\ldots,N; t = 1,\ldots,T,$$

mit dem individuenspezifischen Effekt α_i und ε_{it} wie oben (vgl. S. 12) definiert. Gegebenfalls zusätzlich vorhandene, zeitspezifische Effekte λ_t - aufgrund der Annahme $u_{it} := \alpha_i + \lambda_t + \varepsilon_{it}$ - werden als feste Effekte modelliert und in den Vektor der Strukturparameter aufgenommen (der Vektor $\boldsymbol{x_{it}}$ und somit die Matrix $\boldsymbol{X_i}$ werden entsprechend um Dummy-Variablen erweitert), was bei einem typischen Panel (mit großem N und kleinem T) keine Probleme bereitet.

α_i und ε_{it} seien normalverteilt und für $i, j = 1,\ldots,N;\ t = 1,\ldots,T$ gelte:

$$E(\alpha_i) = 0, \quad E(\varepsilon_{it}) = 0; \quad E(\alpha_i \varepsilon_{jt}) = 0 \quad \forall i,j,t; \quad E(\alpha_i \boldsymbol{x_{it}'}) = \boldsymbol{0}, \quad E(\varepsilon_{it}\boldsymbol{x_{it}'}) = \boldsymbol{0},$$

sowie

$$E(\alpha_i \alpha_j) = \begin{cases} \sigma_\alpha^2 & \text{falls } i = j \\ 0 & sonst \end{cases} \quad \text{und} \quad E(\varepsilon_{it}\varepsilon_{js}) = \begin{cases} \sigma_\varepsilon^2 & \text{falls } i = j \wedge t = s \\ 0 & \text{sonst.} \end{cases}$$

Aus Identifikationsgründen (man kann nur $\frac{\boldsymbol{\beta}}{\sigma_\varepsilon}$ identifizieren, vgl. Ronning, 1991, S. 10) wird zusätzlich unterstellt, daß $\sigma_\varepsilon^2 = 1$.

In diesem nichtlinearen Panelmodell mit zufälligen Effekten gilt demzufolge $E(u_{it}u_{is}) = E(\alpha_i^2) = \sigma_\alpha^2 \neq 0$ für $t \neq s$, mit anderen Worten, für ein bestimmtes Individuum wird eine konstante Korrelation über die Zeit (Äquikorrelation) unterstellt; gleichzeitig wird unterstellt, daß die Entscheidungen verschiedener Individuen unabhängig voneinander sind. Außerdem gilt $E(u_{it}u_{is}) = \sigma_\alpha^2 + \sigma_\varepsilon^2$ für $i = j \wedge t = s$. Damit ist der Störgrößenvektor $\boldsymbol{u_i} = (u_{i1},\ldots,u_{iT})'$ multivariat normalverteilt mit Erwartungswert $\boldsymbol{0}$ und Kovarianzmatrix $\boldsymbol{\Omega}$, mit:

$$\boldsymbol{\Omega} = E(\boldsymbol{u_i}\,\boldsymbol{u_i'}) = \begin{pmatrix} \sigma_\alpha^2+\sigma_\varepsilon^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2+\sigma_\varepsilon^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2+\sigma_\varepsilon^2 \end{pmatrix} = \sigma_\varepsilon^2 \cdot \boldsymbol{I_T} + \sigma_\alpha^2 \cdot \iota_T \iota_T',$$

mit dem T-dimensionalen Einsenvektor $\iota_T = (1,\ldots,1)'$.

Im Fall des hier vorliegenden Probitmodells enthält dann die zugehörige logarithmierte Likelihoodfunktion ein T-faches Normalverteilungsintegral je Individuum.
Um die Berechnung T-dimensionaler Integrale bei der Auswertung der logarithmierten Likelihoodfunktion zu vermeiden, geht man folgendermaßen vor:

Man bestimmt zunächst die bedingte Likelihood für $\boldsymbol{y_i}$ bei gegebenem α_i, da in diesem Fall die y_{it}^* unabhängig normalverteilt sind. Indem diese bedingte Likelihood mit der Dichtefunktion für α_i multipliziert und dann bezüglich α_i integriert wird, erhält man die marginale (unbedingte) Likelihood für $\boldsymbol{y_i}$ (die nicht mehr von den individuenspezifischen Effekten abhängt). Dieses Vorgehen sieht formal folgendermaßen aus:

Ausgehend von der allgemeinen Form der Likelihoodfunktion für die Beobachtungen y_{it}, wie oben spezifiziert (vgl. S. 11), erhält man

$$P(\boldsymbol{y_i}|\boldsymbol{\alpha}_i) = \prod_{t=1}^{T} \boldsymbol{\Phi}(\boldsymbol{x_{it}'\beta} + \alpha_i)^{y_{it}} \left[1 - \boldsymbol{\Phi}(\boldsymbol{x_{it}'\beta} + \alpha_i)\right]^{1-y_{it}},$$

die bezüglich α_i bedingte Likelihood für $\boldsymbol{y_i} = (y_{i1},\ldots,y_{iT})'$, mit $\boldsymbol{\Phi}$, der Verteilungsfunktion der Standardnormalverteilung, wobei für $\boldsymbol{\Phi}(x)$ allgemein gilt

$$\boldsymbol{\Phi}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, exp\left(\frac{-z^2}{2}\right) dz\,.$$

Die Symmetrie der (Standard-) Normalverteilung (es ist: $\boldsymbol{\Phi}(-\boldsymbol{x}) = 1 - \boldsymbol{\Phi}(\boldsymbol{x})$) impliziert, daß (vgl. Greene, 1997, S. 882):

$$\boldsymbol{\Phi}(z)^{y_{it}} \cdot \boldsymbol{\Phi}(z)^{1-y_{it}} = \boldsymbol{\Phi}((2 \cdot y_{it} - 1)\, z)$$

für einen beliebigen Skalar z.

Dies läßt die folgende Vereinfachung der obigen Schreibweise für die bedingte Likelihood zu:

$$P(\boldsymbol{y_i}|\alpha_i) = \prod_{t=1}^{T} \boldsymbol{\Phi}\big((2 \cdot y_{it} - 1)(\boldsymbol{x_{it}'\beta} + \alpha_i)\big)\,.$$

Für die als normalverteilt angenommenen α_i wurde für alle $i = 1,\ldots,N$ unterstellt, daß $E(\alpha_i) = 0$; damit ist nur die Varianz von α_i, σ_α^2, unbekannt.

Um die Abhängigkeit der marginalen Likelihood von σ_α^2 deutlich zu machen, kann man α_i durch $\sigma_\alpha \tilde{\alpha}_i$ ersetzen, so daß $\tilde{\alpha}_i = \dfrac{\alpha_i}{\sigma_\alpha}$ standardnormalverteilt ist.

Die marginale Likelihood für $\boldsymbol{y_i}$ ergibt sich somit zu:

$$P(\boldsymbol{y_i}) = \int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x_{it}'\beta} + \sigma_\alpha \tilde{\alpha}_i\big)\Big)\, \varphi(\tilde{\alpha}_i)\, d\tilde{\alpha}_i \ ,$$

mit $\varphi(\tilde{\alpha}_i)$, der der Verteilungsfunktion $\boldsymbol{\Phi}$ entsprechenden Dichtefunktion. Dies ist richtig, denn ganz allgemein gilt für eine bzgl. x bedingte Verteilungsfunktion von y (vgl. Fisz, 1962, S. 47):

$$\int_{-\infty}^{\infty} F\big(y \,|\, x\big) \cdot f(x)\, dx = F(y) \, .$$

Die marginale Likelihood für $\boldsymbol{y_i}$ hängt damit nur noch vom Strukturparametervektor $\boldsymbol{\beta}$ und der Varianz der zufälligen Effekte, σ_α^2, und nicht nicht mehr von den zufälligen Effekten selbst ab.

Unter Verwendung der Annahme, daß die Zufallsvektoren $\boldsymbol{y_i^*}$ für verschiedene Individuen i unabhängig voneinander sind, gelangt man zu folgender logarithmierter Likelihoodfunktion für die $\boldsymbol{y_i}, i = 1,\ldots,N$:

$$ln\, L = \sum_{i=1}^{N} ln\ P(\boldsymbol{y_i}) = \sum_{i=1}^{N} ln \int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x_{it}'\beta} + \sigma_\alpha \tilde{\alpha}_i\big)\Big)\, \varphi(\tilde{\alpha}_i)\, d\tilde{\alpha}_i \, .$$

Die ML-Schätzung liefert konsistente Schätzwerte für $\boldsymbol{\beta}$ und σ_α^2. Man muß dazu nur (getrennt für jedes Individuum) das Integral über das Produkt von T Standardnormalverteilungsfunktionen lösen. Mit Hilfe der Gauß-Hermite-Quadratur (siehe Abschnitt 4.3) können diese Integrale aber numerisch effizient gelöst werden (vgl. Butler und Moffitt, 1982, S. 762).

Im Vergleich zu dem Modell mit festen Effekten erkennt man, daß hier das Problem einer mit N wachsenden Anzahl individuenspezifischer Parameter, α_i, entfällt. Dafür werden allerdings restriktive Verteilungsannahmen für die individuenspezifischen Effekte getroffen, d.h., trifft die Annahme nicht zu, daß die α_i normalverteilt sind, so führt dies im allgemeinen zu inkonsistenten Parameterschätzwerten für $\boldsymbol{\beta}$ und σ_α^2.

Abschließend soll in diesem Abschnitt noch einmal darauf hingewiesen werden, daß das hier behandelte statische Probitmodell mit stochastischen Effekten zwar eine intertemporale Korrelation der Störgrößen bzw. der latenten Zufallsvariablen aufweist, diese jedoch über die Zeit als konstant angenommen wird. Für größere T, für die diese Annahme sich allgemein als unzutreffend erweisen dürfte, ist die obige Spezifikation deshalb meist unbefriedigend. So erscheint es zumindest nicht unplausibel, eine über die Zeit abnehmende Korrelation der Störgrößen zu unterstellen. Dieser Aspekt wird in Abschnitt 3.2.1 aufgegriffen.

Eine weitere restriktive Annahme dieses Modells ist die Annahme der Unkorreliertheit von α_i und $\boldsymbol{x_{it}}$. Diese Annahme wird im Chamberlain-Modell aufgegeben.

2.2.2.2 Das statische Probitmodell für korrelierte α_i und $\boldsymbol{x_{it}}$ - Der Chamberlain-Ansatz

Bei dem im vorangehenden Abschnitt behandelten Probitmodell mit zufälligen Effekten (und damit äquikorrelierten Störgrößen) wurde angenommen, daß die stochastischen individuenspezifischen Effekte (α_i) nicht mit dem Vektor der exogenen Variablen $\boldsymbol{x_{it}}$ korreliert sind. Diese Annahme wird im Chamberlain-Modell fallengelassen. Es ist somit die bedingte Verteilung der α_i bei gegebenem $\boldsymbol{x_{it}}$ zu spezifizieren.

Ausgehend von dem o.a. Schwellenwertmodell und der Spezifikation der latenten Variablen durch das Modell

$$y_{it}^* = \boldsymbol{x_{it}'\beta} + \alpha_i + \varepsilon_{it}, \quad i = 1,\ldots,N;\, t = 1,\ldots,T,$$

mit $y_{it}^*, \boldsymbol{x_{it}}, \boldsymbol{\beta}$, $\boldsymbol{\alpha}_i$ und $\boldsymbol{\varepsilon}_{\mathrm{it}}$ wie oben definiert unterstellte Chamberlain (1980, 1985) für die stochastischen, individuenspezifischen Effekte $\boldsymbol{\alpha}_i$ folgenden Zusammenhang:

Es ist $\boldsymbol{\alpha}_i = \boldsymbol{a'x_i} + v_i$, mit $\boldsymbol{a'} = (a_{11},\ldots,a_{1K},a_{21},\ldots,a_{2K},\ldots,a_{T1},\ldots,a_{TK})$ einem Zeilenvektor der Dimension $(T \cdot K)$, $\boldsymbol{x_i'} = (x_{i11},\ldots,x_{i1K},x_{i21},\ldots,x_{i2K},\ldots,x_{iT1},\ldots,x_{iTK})$, dem $(T \cdot K)$-dimensionalen Vektor der exogenen Variablen für das Individuum i und v_i einer unabhängig und identisch (normal-) verteilten Störgröße mit Erwartungswert null und Varianz σ_v^2.
Diese Spezifikation unterstellt, daß der bedingte Erwartungswert von α_i für gegebenes $\boldsymbol{x_i}$ linear ist und α_i einer bestimmten Wahrscheinlichkeitsverteilung genügt. Außerdem geht Chamberlain davon aus, daß die Störgröße v_i von $\boldsymbol{x_i}$ unabhängig ist.

Durch Substitution von $\alpha_i = \boldsymbol{a'x_i} + v_i$ in $y_{it}^* = \boldsymbol{x_{it}'\beta} + \alpha_i + \varepsilon_{it}$ erhält man:

$$y_{it}^* = \boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i} + v_i + \varepsilon_{it}, \quad i = 1,\ldots,N;\, t = 1,\ldots,T.$$

Unter der Annahme, daß die Störgröße ε_{it} stochastisch unabhängig und identisch standardnormalverteilt ist, gelangt man zu einer logarithmierten Likelihoodfunktion, analog der des Probitmodells mit zufälligen Effekten:

$$ln\, L = \sum_{i=1}^{N} ln \int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\big((2 \cdot y_{it} - 1)(\boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i} + \sigma_v \tilde{v}_i)\big)\, \varphi(\tilde{v}_i)\, d\tilde{v}_i\,.$$

Dabei bezeichnet $\boldsymbol{\Phi}$ auch hier die Verteilungsfunktion der Standardnormalverteilung, φ die zugehörige Dichtefunktion, und für $\tilde{v}_i$ gilt, analog $\tilde{\alpha}_i$, $\tilde{v}_i = \dfrac{v_i}{\sigma_v}$.

Der grundlegende Unterschied zum Probitmodell mit äquikorrelierten Störgrößen liegt demnach in der Einbeziehung des Terms $\boldsymbol{a'x_i}$ in die logarithmierte Likelihoodfunktion, womit die Abhängigkeit zwischen den individuenspezifischen Parametern α_i und sämtlichen exogenen Variablen für das Individuum i erfaßt wird.

Die eigentlich notwendige Lösung des obigen Integrals umgeht man dadurch, daß man T voneinander getrennte Parameterschätzungen vornimmt. Für jede der T Wellen liegt dann ein univariates Probitmodell vor, und die bezüglich α_i marginale (unbedingte) Wahrscheinlichkeit für y_{it}, kann man wie folgt schreiben:

$$P(y_{it} = 1) = \boldsymbol{\Phi}\left(\frac{1}{\sqrt{1+\sigma_v^2}}\left(\boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i}\right)\right),$$

denn für ein beliebiges, festes $t, t = 1,\ldots,T$ ergibt sich aufgrund der Symmetrieeigenschaft der Verteilung folgender Zusammenhang:

$$P(y_{it} = 1) = P(y_{it}^* > 0) = P(\boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i} + v_i + \varepsilon_{it} > 0) = P(v_i + \boldsymbol{\varepsilon}_{it} > -(\boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i}))$$

$$= P(\, v_i + \varepsilon_{tt} < \boldsymbol{x_{it}'\beta} + \boldsymbol{a'x_i}\,),$$

wobei die Störgröße $v_i + \varepsilon_{it}$ normalverteilt ist, mit Erwartungswert null und Varianz $1 + \sigma_v^2$.

Für die auf diese Art ermittelten ML-Schätzer gilt:

Faßt man die latenten Variablen y_{it}^* für alle $t = 1,\ldots,T$ in einem Vektor zusammen, dann erhält man

$$\boldsymbol{y_i^*} = (\boldsymbol{I_T} \otimes \boldsymbol{\beta'} + \boldsymbol{\iota}_T \boldsymbol{a'})\, \boldsymbol{x_i} + \boldsymbol{\iota}_T v_i + \varepsilon_i = \boldsymbol{\Pi}\, \boldsymbol{x_i} + \boldsymbol{\iota}_T v_i + \varepsilon_i, \quad i = 1,\ldots,N\,.$$

Es ist $\boldsymbol{I_T}$ eine Einheitsmatrix der Ordnung T und $\boldsymbol{\iota}_T$ ein T-dimensionaler Vektor, der ausschließlich Einsen enthält, und es ist $\boldsymbol{I_T} \otimes \boldsymbol{\beta'}$ das Kronecker-Produkt von $\boldsymbol{I_T}$ und $\boldsymbol{\beta'}$.

Vektorisiert man nun $\boldsymbol{\Pi}$, so daß

$$\boldsymbol{\pi}' = \left[vec(\boldsymbol{\Pi})\right]' = \frac{1}{\sqrt{1+\boldsymbol{\sigma}_v^2}}(a_{11}+\boldsymbol{\beta}_1,\ldots,a_{1K}+\boldsymbol{\beta}_K,a_{21},\ldots,a_{2K},\ldots,a_{T1},\ldots,a_{TK};a_{11},\ldots,a_{1K},a_{21}+\boldsymbol{\beta}_1$$
$$,\ldots,a_{2K}+\boldsymbol{\beta}_K,\ldots,a_{T1},\ldots,a_{TK};\ldots;a_{11},\ldots,a_{1K},\ldots,a_{T1}+\boldsymbol{\beta}_1,\ldots,a_{TK}+\boldsymbol{\beta}_K)$$
$$= (\boldsymbol{\pi}_1';\boldsymbol{\pi}_2';\ldots;\boldsymbol{\pi}_K')$$

dann erkennt man, daß für $\hat{\boldsymbol{\pi}}_t$ für alle $t = 1,\ldots,T$ folgende Restriktionen gelten:

$$\hat{\boldsymbol{\pi}}_t = \frac{1}{\sqrt{1+\hat{\boldsymbol{\sigma}}_v^2}}(\hat{a}_{11},\ldots,\hat{a}_{1K},\ldots,\hat{a}_{t1}+\hat{\boldsymbol{\beta}}_1,\ldots,\hat{a}_{tK}+\hat{\boldsymbol{\beta}}_K,\ldots,\hat{a}_{T1},\ldots,\hat{a}_{TK}).$$

Die Elemente von $\boldsymbol{\pi}$ schätzt man z.B. mit einem ′Minimum-distance′-Ansatz (vgl. Chamberlain, 1984, S. 1288ff.), unter Einbeziehung der jeweiligen Restriktionen. Dieses sieht im vorliegenden Fall folgendermaßen aus:

Sei $\boldsymbol{\theta} = (\boldsymbol{\beta}',\mathbf{a}')'$, dann ist

$$\boldsymbol{\pi} = f(\boldsymbol{\theta})$$

die allgemeine Form der $\boldsymbol{\pi}$ auferlegten Restriktionen und die Minimierung des Ausdrucks

$$(\hat{\boldsymbol{\pi}} - f(\boldsymbol{\theta}))'\hat{\boldsymbol{\Omega}}^{-1}(\hat{\boldsymbol{\pi}} - f(\boldsymbol{\theta}))$$

nach $\boldsymbol{\theta}$, liefert den ′Minimum-distance′-Schätzer für $\boldsymbol{\theta}$. Dabei ist $\hat{\boldsymbol{\Omega}}$ der Schätzer der asymptotischen Kovarianzmatrix von $\hat{\pi}$ (vgl. Chamberlain, 1984, S. 1288ff.).

In diesem und dem vorangegangenen Abschnitt wurden die logarithmierten Likelihood-funktionen wiedergegeben, welche zur ML-Schätzung der dort betrachteten Modelle benötigt werden. Auf die ML-Schätzung selbst wurde dabei nicht eingegangen. Abschnitt 2.3 liefert nun Hinweise auf die ML-Schätzung in binomialen Antwortmodellen, wie sie auch den Querschnittsanalysen zugrunde liegen. Weitere Hinweise zur Implementierung statischer und dynamischer Panelmodelle werden in Kapitel 4 behandelt.

2.3 ML-Schätzung in binomialen Antwortmodellen

Eine der fundamentalen Feststellungen im Zusammenhang mit der Schätzung von Probit- und Logitmodellen ist, daß sämtliche partiellen ersten Ableitungen der jeweiligen logarithmierten Likelihoodfunktion nach den K Elementen des Parametervektors $\boldsymbol{\beta}$ nichtlinear in $\boldsymbol{\beta}$ sind. Deshalb benötigt man in diesen Fällen iterative Methoden, um eine ML-Schätzung durch-führen zu können. In den hier betrachteten binomialen Probitmodellen wird deshalb zur nicht-

linearen Optimierung der logarithmierten Likelihoodfunktion die Newton-Raphson-Methode verwendet.

Abschnitt 2.3.1 faßt die wesentlichen Merkmale dieser Methode zusammen; dazu werden zunächst die partiellen ersten und zweiten Ableitungen der logarithmierten Likelihood-funktion nach dem Parametervektor $\boldsymbol{\beta}$ wiedergegeben.

Abschnitt 2.3.2 zeigt dann verschiedene Schätzer für die asymptotische Kovarianzmatrix des ML-Schätzers von $\boldsymbol{\beta}$, wobei der Schwerpunkt auf der von Berndt, Hall, Hall und Hausman entwickelten sogenannten BHHH-Methode (vgl. Berndt, Hall, Hall and Hausman, 1974, S. 653ff.) liegt. Ein derartiger Schätzer für die asymptotische Kovarianzmatrix des ML-Schätzers von $\boldsymbol{\beta}$ wird benötigt, um u.a. Hypothesentests über die K Elemente dieses Parametervektors durchführen zu können.

Im anschließenden Abschnitt 2.3.3 wird eine im Rahmen der ML-Schätzung häufig verwendete Testprozedur, der Likelihood-Quotienten-Test dargestellt.

2.3.1 Grundzüge der Newton-Raphson-Methode

Zur Bestimmung des ML-Schätzers $\hat{\boldsymbol{\beta}}$ mittels der Newton-Raphson-Methode werden die partiellen ersten und zweiten Ableitungen der logarithmierten Likelihoodfunktion nach dem Parametervektor $\boldsymbol{\beta}$ benötigt.

Ausgehend von dem Schwellenwertmodell

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst,} \end{cases}$$

mit $y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T$

und mit $y_{it}^*, \boldsymbol{x}_{it}, \boldsymbol{\beta}$ und u_{it}, wie oben definiert, und mit u_{it} stochastisch unabhängig und identisch über die N Individuen und die T Wellen verteilt, erhält man für ein binomiales Antwortmodell mit $N \cdot T$ Beobachtungen:

Die Wahrscheinlichkeit, daß das Individuum i in der Welle t den Wert eins annimmt ergibt sich zu (vgl. Abschnitt 2.2):

$$P(y_{it} = 1) = F(\boldsymbol{x}_{it}'\boldsymbol{\beta}) .$$

Unterstellt man für die Verteilungsfunktion F der Störgrößen u_{it} eine Standardnormalverteilung so führt dies zu einem Probitmodell.

Für das Probitmodell gilt dann (vgl. Greene, 1993, S. 637):

$$P(y_{it}=1)=\int_{-\infty}^{\boldsymbol{x}_{it}'\boldsymbol{\beta}} \boldsymbol{\varphi}(t)\, dt = \boldsymbol{\Phi}(\boldsymbol{x}_{it}'\boldsymbol{\beta}) \ ,$$

mit φ der Dichtefunktion, und $\boldsymbol{\Phi}$, der Verteilungsfunktion der Standardnormalverteilung.

Ziel einer Schätzung ist die Ermittlung von Schätzwerten insbesondere für $\boldsymbol{\beta}$. Es ist aber wichtig festzuhalten, daß in diesem Antwortmodell, wie in jeder nichtlinearen Regression, die Modellparameter $\beta_1,\ldots,\beta_K$ nicht notwendigerweise die marginalen Effekte wiedergeben, denn es gilt ganz allgemein (vgl. Greene, 1993, S. 639):

$$\frac{\partial\, E(y)}{\partial\, \boldsymbol{x}} = \left\{ \frac{\partial\, F(\boldsymbol{x}'\boldsymbol{\beta})}{\partial\, (\boldsymbol{x}'\boldsymbol{\beta})} \right\} \cdot \boldsymbol{\beta} = f(\boldsymbol{x}'\boldsymbol{\beta}) \cdot \boldsymbol{\beta} \ ,$$

wobei f die zu F gehörige Dichtefunktion und E den Erwartungswertoperator bezeichnen. Die marginalen Effekte sind deshalb von $\boldsymbol{x}$ abhängig.

Dementsprechend erhält man für die marginalen Effekte des Probitmodells:

$$\frac{\partial\, E(y_{it})}{\partial\, \boldsymbol{x}_{it}} = \boldsymbol{\varphi}(\boldsymbol{x}_{it}'\boldsymbol{\beta}) \cdot \boldsymbol{\beta} \, .$$

Unter der Annahme der Unabhängigkeit der Störgrößen u_{it} ergibt sich folgende allgemeine Likelihoodfunktion für die Beobachtungen $y_{it}, i = 1,\ldots,N;\ t = 1,\ldots,T$ eines nichtlinearen Modells für Paneldaten (vgl. S. 11):

$$L = \prod_{i=1}^{N} \prod_{t=1}^{T} \left[F(\boldsymbol{x}_{it}'\boldsymbol{\beta}) \right]^{y_{it}} \left[1 - F(\boldsymbol{x}_{it}'\boldsymbol{\beta}) \right]^{1-y_{it}}$$

und die logarithmierte Likelihoodfunktion für eine Stichprobe von N Individuen über T Wellen ist demzufolge:

$$ln\, L = \sum_{i=1}^{N} \sum_{t=1}^{T} (y_{it}\, ln\, F_{it} + (1 - y_{it})\, ln\, (1 - F_{it})) \, ,$$

mit $F_{it} := F(\boldsymbol{x}_{it}'\boldsymbol{\beta})$.

Die notwendigen Bedingungen für ein Maximum dieser Funktion fordern, daß

$$\boldsymbol{s}(\boldsymbol{\beta}) := \frac{\partial\ ln\, L}{\partial\, \boldsymbol{\beta}} = \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\frac{f_{it}}{F_{it}\,(1 - F_{it})} \right] (y_{it} - F_{it})\, \boldsymbol{x}_{it} = \boldsymbol{0} \, ,$$

mit $f_{it} = f(\boldsymbol{x}_{it}'\boldsymbol{\beta})$ der zur jeweiligen Verteilungsfunktion F_{it} gehörigen Dichtefunktion (vgl. Amemiya, 1981, S. 1495).

Die partiellen ersten Ableitungen der logarithmierten Likelihoodfunktion nach $\boldsymbol{\beta}$ sind folglich nichtlinear in $\boldsymbol{\beta}$ und für die Lösung dieses Schätzproblems ist deshalb die Verwendung einer iterativen Methode, wie z.B. der Newton-Raphson-Methode, notwendig.
Zur Durchführung dieses Verfahrens und zugleich zur Herleitung eines Schätzers für die asymptotische Kovarianzmatrix der Parameterschätzfunktion benötigt man zusätzlich zu dem Gradienten der ML-Funktion die Matrix der partiellen zweiten Ableitungen der logarithmierten Likelihoodfunktion, die Hesse-Matrix $\boldsymbol{H}$. Für das binomiale Antwortmodell mit $N \cdot T$ Beobachtungen erhält man:

$$\boldsymbol{H} = \frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}'} = -\sum_{i=1}^{N}\sum_{t=1}^{T} \left[\frac{y_{it} - F_{it}}{F_{it}(1-F_{it})}\right]^2 (f_{it})^2\, \boldsymbol{x_{it}}\, \boldsymbol{x'_{it}} + \sum_{i=1}^{N}\sum_{t=1}^{T} \frac{y_{it} - F_{it}}{F_{it}(1-F_{it})} f'_{it}\; \boldsymbol{x_{it}} \boldsymbol{x'_{it}},$$

mit f'_{it}, der partiellen ersten Ableitung der Dichtefunktion f_{it} (vgl. Amemiya, 1981, S. 1495).

Die Newton-Raphson-Methode läßt sich - für die *k*-te Iteration - nun wie folgt schreiben:

$$\hat{\boldsymbol{\beta}}_k = \hat{\boldsymbol{\beta}}_{k-1} - [\boldsymbol{H}]^{-1}_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_{k-1}}\, \boldsymbol{s}(\hat{\boldsymbol{\beta}}_{k-1}),$$

mit $\hat{\boldsymbol{\beta}}_k$, dem Wert von $\hat{\boldsymbol{\beta}}$ bei der *k*-ten Iteration.

Für die zu betrachtenden Probitmodelle ist die logarithmierte Likelihoodfunktion global konkav, da die Hesse-Matrix negativ definit ist. In diesem Fall führt die Newton-Raphson-Methode - unabhängig vom Startwert - zu einer eindeutigen Lösung, und diese Lösung ist ein Maximum (vgl. Amemiya, 1985, S. 273f.).

Grundlage der Newton-Raphson-Methode ist eine lineare Taylorreihenentwicklung der Form

$$\frac{\partial F(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \cong g(\boldsymbol{\theta}_0) + \boldsymbol{H}(\boldsymbol{\theta}_0) \cdot (\boldsymbol{\theta} - \boldsymbol{\theta}_0) = \boldsymbol{0}$$

mit $F(\boldsymbol{\theta})$, der bzgl. $\boldsymbol{\theta}$ zu maximierenden Funktion, für die der Gradient $\frac{\delta F(\boldsymbol{\theta})}{\delta \boldsymbol{\theta}} = g(\boldsymbol{\theta})$ mittels einer linearen Taylorreihenentwicklung um einen beliebigen Vektor $\boldsymbol{\theta}_0$ approximiert wird (vgl. Greene, 1997, S. 203).
Dabei bezeichnet $\boldsymbol{H}(\boldsymbol{\theta}_0)$ die Matrix der partiellen zweiten Ableitungen von $F(\boldsymbol{\theta})$ nach dem Vektor $\boldsymbol{\theta}$, ausgewertet an der Stelle $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ und $g(\boldsymbol{\theta}_0)$ den Gradienten der Funktion $F(\boldsymbol{\theta})$, ausgewertet an der Stelle $\boldsymbol{\theta} = \boldsymbol{\theta}_0$.

Durch Auflösen der obigen Vektorgleichung nach $\boldsymbol{\theta}$ und Gleichsetzen von $\boldsymbol{\theta}$ mit $\boldsymbol{\theta}_{t+1}$ und von $\boldsymbol{\theta}_0$ mit $\boldsymbol{\theta}_t$ erhält man die der Newton-Raphson-Methode zugrundeliegende Iterationsgleichung

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \boldsymbol{H}^{-1}(\boldsymbol{\theta}_t)\, g(\boldsymbol{\theta}_t)\,.$$

Unter bestimmten Regularitätsbedingungen (vgl. Amemiya, 1985, S. 270) ist der ML-Schätzer $\hat{\boldsymbol{\beta}}$ konsistent, asymptotisch normalverteilt und asymptotisch effizient. Diese wünschenswerten asymptotischen Eigenschaften machen die Attraktivität des ML-Schätzers aus.

Die Ermittlung der Hesse-Matrix der logarithmierten Likelihood-Funktion dient jedoch nicht allein der Herleitung des ML-Schätzers für $\boldsymbol{\beta}$, sondern auch der Herleitung für die asymptotische Kovarianzmatrix des ML-Schätzers.

2.3.2 Schätzung der asymptotischen Kovarianzmatrix des ML-Schätzers

Die asymptotische Kovarianzmatrix $\boldsymbol{V}(\hat{\boldsymbol{\beta}})$ des ML-Schätzers kann auf verschiedene Arten hergeleitet werden. Ein möglicher Schätzer geht aus von der Inversen der Informationsmatrix $Inf(\boldsymbol{\beta})$ ausgewertet an der Stelle des ML-Schätzers $\hat{\boldsymbol{\beta}}$:

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \left[Inf(\boldsymbol{\beta})\right]^{-1}_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \quad ,$$

wobei $Inf(\boldsymbol{\beta})$ definiert ist als (-1)-fache des Erwartungswertes der Hesse-Matrix:

$$Inf(\boldsymbol{\beta}) = -E\left(\frac{\partial^2 \ln L}{\partial \boldsymbol{\beta}\, \partial \boldsymbol{\beta}'}\right) = -E(\boldsymbol{H})\,.$$

Die der Informationsmatrix zugrundeliegende Hesse-Matrix $\boldsymbol{H}$ enthält jedoch fast immer nichtlineare Funktionen der Daten, deren Erwartungswert im allgemeinen unbekannt ist (vgl. Greene, 1997, S. 138f.). Damit ist in diesem Fall die Schätzung der asymptotischen Kovarianzmatrix auf der Basis der Informationsmatrix nicht zu realisieren.

Eine mögliche Alternative liefert die Inverse der negativen Hesse-Matrix ausgewertet an der Stelle des ML-Schätzers:

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = -[\boldsymbol{H}]^{-1}_{\boldsymbol{\beta}=\hat{\boldsymbol{\beta}}} \quad .$$

Die Herleitung der partiellen zweiten Ableitungen der logarithmierten Likelihoodfunktion nach dem Parametervektor $\boldsymbol{\beta}$, d.h., die Herleitung der Hesse-Matrix $\boldsymbol{H}$, erweist sich jedoch häufig als äußerst aufwendig.
Deshalb sind solche Methoden vorteilhaft, für die nur die partiellen ersten Ableitungen benötigt werden. Eine solche Methode ist die von Berndt, Hall, Hall und Hausman ent-wickelte BHHH-Methode. Diese Methode basiert auf dem Ergebnis, daß der Erwartungswert der

Hesse-Matrix gleich ist der Kovarianzmatrix des Gradientenvektors (vgl. Greene, 1997, S. 139).

Als Schätzer $\boldsymbol{V}(\hat{\boldsymbol{\beta}})$ der asymptotischen Kovarianzmatrix des Strukturparametervektors $\hat{\boldsymbol{\beta}}$ ergibt sich hier folgende Matrix (vgl. Greene, 1997, S.139):

$$\boldsymbol{V}(\hat{\boldsymbol{\beta}}) = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} \hat{g}_{it} \hat{g}_{it}' \right]^{-1},$$

wobei $\hat{g}_{it} = \dfrac{\partial \ln f(\boldsymbol{x_{it}}, \hat{\boldsymbol{\beta}})}{\partial \hat{\boldsymbol{\beta}}}$.

Keiner der drei angesprochenen Schätzer ist einem anderen aufgrund seiner statistischen Eigenschaften vorzuziehen; alle drei Schätzer sind asymptotisch äquivalent. Die Schätzer können jedoch, angewandt auf endliche Stichproben, unterschiedliche Schätzwerte liefern, diese Tatsache kann zu unterschiedlichen statistischen Schlußfolgerungen (z.B. bei der Beurteilung der Signifikanz einzelner Parameter) führen. Die Bestimmung des BHHH-Schätzers erfordert allerdings in den meisten Fällen den vergleichbar geringsten Rechenaufwand (soweit die Bestimmung der jeweiligen Schätzer faktisch möglich ist), woraus sich eine gewisse Vorteilhaftigkeit dieser Vorgehensweise herleiten läßt.

2.3.3 Zur Verwendung des Likelihood-Quotienten-Tests

Zur Überprüfung allgemeiner Hypothesen bzgl. β wie z.B. von Nullhypothesen der Form H_0: $c(\boldsymbol{\beta}) = 0$, existieren drei asymptotisch äquivalente Testverfahren; der Likelihood-Quotienten-Test (LR-Test), der Wald-Test sowie der Lagrange-Multiplikator-Test (vgl. Darnell, 1994, S. 267ff.).

Der LR-Test testet die Differenz aus dem Wert der logarithmierten Likelihoodfunktion an der Stelle des unrestringierten ML-Schätzers und dem Wert der logarithmierten Likelihoodfunktion an der Stelle des restringierten ML-Schätzers (ML-Schätzer, der die Restriktion $c(\boldsymbol{\beta}) = 0$ erfüllt). Der Wald-Test beruht dagegen auf der Tatsache, daß die Restriktion, angewendet auf den ML-Schätzer, einen Wert nahe null liefern sollte ($c(\hat{\boldsymbol{\beta}}) \approx 0$), falls sie gültig ist. Beim Lagrange-Multiplikator-Test erwartet man für die Ableitung der logarithmierten Likelihoodfunktion an der Stelle des restringierten ML-Schätzers einen Wert nahe null, falls die Restriktion gültig ist.

Die drei Testverfahren sind zwar asymptotisch äquivalent, können jedoch in kleinen Stichproben zu unterschiedlichen Entscheidungen führen. Die Eigenschaften dieser Teststatistiken für kleine Stichproben sind dabei, abgesehen von Sonderfällen, unbekannt. Welches der drei

Testverfahren man anwendet, hängt deshalb typischerweise davon ab, welche Teststatistik einfacher zu berechnen ist.

In der weiteren Arbeit findet nur der Likelihood-Quotienten-Test Verwendung; dieser wird deshalb im folgenden genauer betrachtet:

Der Test der Nullhypothese

$$H_0\colon c(\boldsymbol{\beta}) = 0$$

basiert beim Likelihood-Quotienten-Test auf dem Likelihood-Quotienten λ, der definiert ist, als der Quotient aus dem Wert der Likelihood an der Stelle des restringierten ML-Schätzers ($\hat{L}_R$) und dem Wert der Likelihood an der Stelle des unrestringierten ML-Schätzers ($\hat{L}_U$). Der Wert des Parameters λ liegt damit zwischen null und eins.

Die Likelihood-Quotienten-Teststatistik (LR) ist definiert als

$$LR = -2 \cdot ln\,\lambda = 2 \cdot (ln\,\hat{L}_U - ln\,\hat{L}_R)$$

und unter $c(\boldsymbol{\beta}) = 0$ asymptotisch χ_q^2-verteilt, mit q Freiheitsgraden. Es ist q die Anzahl der in der Nullhypothese enthaltenen Restriktionen. Die Nullhypothese wird verworfen, falls der Wert der LR-Teststatistik den Wert des entsprechenden Quantils der χ_q^2-Verteilung überschreitet.

Nachdem hiermit verschiedene Aspekte der ML-Schätzung in binomialen Antwortmodellen betrachtet wurden, geht es im Kapitel 3 um die Spezifikation dynamischer nichtlinearer Panelmodelle.

3 Dynamische nichtlineare Panelmodelle

3.1 Vorbemerkungen

In diesem Abschnitt werden dynamische Probitmodelle auf der Grundlage eines sehr allgemeinen Modells von Heckman (1981a) betrachtet. Man verallgemeinert die bisher betrachteten Modelle, indem man einerseits verzögert endogene Variablen als erklärende Variablen in das jeweilige Schwellenwertmodell einbezieht und andererseits komplexere Störgrößenprozesse als die in den vorangegangenen Abschnitten diskutierten Varianzkomponentenmodelle zuläßt, zwei Varianten, die in der empirischen Analyse die Realitätsnähe des Modells erhöhen können.

Durch das Einbeziehen verzögert endogener Variablen kann man beispielsweise den Fall berücksichtigen, daß die Arbeitslosigkeit einer betrachteten Person in der Vergangenheit einen Einfluß auf den momentanen und/oder zukünftigen Arbeitsstatus dieser Person besitzt. Komplexere Störgrößenprozesse, andererseits, werden einbezogen, wenn das bisher betrachtete Varianzkomponentenmodell zur Abbildung entsprechender ökonomischer Vorgänge nur unzureichend geeignet ist.

Ohne zunächst auf die vielfältigen Schätzprobleme im Zusammenhang mit den, in diesem Abschnitt betrachteten, Modellen einzugehen (vgl. hierzu Abschnitt 2.3 und Kapitel 4), soll aber, um die Notwendigkeit einer derartigen Modellspezifikation für die vorliegende Arbeit deutlich zu machen, bereits an dieser Stelle auf eine Fragestellung hingewiesen werden, welche mit diesen Modellerweiterungen untersucht werden kann und wird:

Bei der Analyse des Arbeitsstatus von Individuen über die Zeit stellt man häufig fest, daß Individuen, welche in der Vergangenheit arbeitslos waren, auch in der Gegenwart oder Zukunft mit größerer Wahrscheinlichkeit nicht beschäftigt sind, als solche Individuen, welche in der Vergangenheit nicht arbeitslos waren. Hierzu werden in der Literatur unterschiedliche Erklärungsansätze diskutiert:

- Der eine Ansatz geht davon aus, daß sich durch den Eintritt eines Ereignisses (Arbeitslosigkeit) die Präferenzen bzw. die Rahmenbedingungen, welche für zukünftige Entscheidungen (eine Tätigkeit aufzunehmen bzw. abzulehnen) bzw. Dispositionen (arbeitslos zu werden) des Individuums relevant sind, verändert haben. Somit läge ein Effekt einer in der Vergangenheit bestehenden Arbeitslosigkeit auf den zukünftigen Arbeitsstatus vor. Diesen Fall bezeichnet Heckman als ′(wahre) Zustandsabhängigkeit′ (′true state dependence′; Heckman, 1981a, S. 115).

– Ein anderer Erklärungsansatz geht von einer vorhandenen (natürlichen) Heterogenität der Individuen aus. Hierbei wird unterstellt, daß die Individuen sich grundsätzlich hinsichtlich ihrer Disposition, arbeitslos zu werden unterscheiden (′unbeobachtete Heterogenität′). Falls diese nicht meßbaren, individuellen Unterschiede über die Zeit hin korreliert sind, und dieser Tatbestand nicht explizit im Modellansatz berücksichtigt wird, kann man aufgrund der Schätzergebnisse zu der falschen Schlußfolgerung gelangen, daß die Zustände (Beschäftigungsstatus) eines Individuums in der Vergangenheit, auch in einem eigentlich statischen Modell, einen Einfluß auf den gegenwärtigen Zustand (Beschäftigungsstatus) dieses Individuums besitzen. Diesen Fall bezeichnet Heckman als ′scheinbare Zustandsabhängigkeit′ (′spurious state dependence′; siehe Heckman, 1981a, S. 115).

Um eine formale Diskriminierung zwischen beiden Erklärungsansätzen mit Hilfe der zu betrachtenden Modelle zu ermöglichen, soll zunächst das Modell von Heckman, welches zur allgemeinen Spezifikation dynamischer nichtlinearer Panelmodelle dient, (in einer etwas vereinfachten Form) dargestellt werden. Danach werden zwei Varianten dieses Modells betrachtet: Zuerst ein statisches Probitmodell mit autokorrelierten Störgrößen (Abschnitt 3.2.1), anschließend ein dynamisches Probitmodell mit äquikorrelierten Störgrößen (Abschnitt 3.2.2). Dabei steht die, hinsichtlich der Modellspezifikation notwendige, Unterscheidung zwischen der Zustandsabhängigkeit und unbeobachteter Heterogenität im Vordergrund. Schließlich sollen diese beiden Modelle kombiniert werden, d.h., im dynamischen Probitmodell sollen zusätzlich autokorrelierte Störgrößen zugelassen werden (Abschnitt 3.2.3).

3.2 Spezifikation dynamischer Probitmodelle

Auch Heckman (1981a) geht bei seiner Spezifikation dynamischer nichtlinearer Panelmodelle davon aus, daß die beobachtbaren binären Zufallsvariablen y_{it} durch latente stetige Zufallsvariablen y_{it}^* erzeugt werden, die einen Schwellenwert überschreiten. Die allgemeine Form des Modells von Heckman lautet (vgl. Heckman, 1981a, S. 121f.):

$$y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \sum_{l=1}^{\infty}\gamma_l y_{i,t-l} + \lambda\sum_{s=1}^{\infty}\prod_{l=1}^{s} y_{i,t-l} + G(L)y_{it}^* + u_{it}\ , \quad i = 1,\ldots,N;\ t = 1,\ldots,T\ .$$

$y_{it}^*, \boldsymbol{x}_{it}', \beta$ und u_{it} seien wie oben (vgl. S. 9) definiert, und Heckman unterstellt, daß die Störgrößen u_{it} unabhängig von $\boldsymbol{x}_{it}$ verteilt sind (für alle i und t). Der Störgrößenvektor $\boldsymbol{u}_i$ wird als multivariat normalverteilt mit Erwartungswertvektor $\boldsymbol{0}$ und beliebiger positiv definiter Kovarianzmatrix $\boldsymbol{\Sigma}$ angenommen.

In diesem Modell mißt der Koeffizient γ_l den Effekt einer Statusänderung, welche mit einer Verzögerung von l Perioden auf die latente Zufallsvariable y_{it}^* wirkt. Im folgenden wird

vereinfachend unterstellt, daß $\gamma_1 = \gamma$ und $\gamma_l = 0$ für $l > 1$ ist, d.h., es wird, soweit das Modell dynamisch ist, von einem Markov-Prozeß erster Ordnung ausgegangen.

Hingegen gibt der Term $\sum_{s=1}^{\infty}\prod_{l=1}^{s} y_{i,t-l}$ an, wie lange sich das Individuum i, ausgehend vom Zeitpunkt t, bereits im ′Zustand 1′ (arbeitslos) befindet. Die Hereinahme dieses Terms, ′Dauer eines Zustands′ (vgl. Hamerle, 1994, S. 1), als exogene Variable in die obige Gleichung ergibt sich aus der theoretischen Überlegung, daß eine Art Akkumulationsprozeß beginnt, sobald ein Individuum einen bestimmten Zustand angenommen hat. Solange beispielsweise ein Individuum i in einem bestimmten Unternehmen beschäftigt ist, baut es Humankapital auf, welches (teilweise) verloren geht, wenn das Individuum diese Firma verläßt. Dieser Term soll aber bei der Spezifikation der dynamischen Modelle in der vorliegenden Arbeit vernachlässigt werden, da er im Rahmen einer Modellierung individueller Arbeitslosigkeit kaum interpretierbar ist.

Bezüglich des Summanden $G(L)y_{it}^*$ gilt folgendes: Ist $G(L)$ ein Lagoperator der Ordnung H, so daß

$$G(L) = g_1 L + g_2 L^2 + ... + g_H L^H \text{ mit } L^H y_{it}^* = y_{i,t-H}^* \text{ und } G(0) = 0,$$

dann verkörpert der Term $G(L)y_{it}^*$ die Idee des Festhaltens an einer Gewohnheit (′habit persistence′; Heckman, 1981a, S. 122) und repräsentiert beispielsweise den Effekt einer früheren Disposition, arbeitslos zu werden, auf die gegenwärtige Disposition nicht beschäftigt zu sein. Dieser Term wird aus Vereinfachungsgründen in den folgenden Betrachtungen vernachlässigt.

Das den folgenden Überlegungen zugrunde liegende Modell ist also

$$y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \gamma\, y_{i,t-1} + u_{it}\,, \quad i = 1,\dots,N;\ t = 1,\dots,T \text{ und}$$

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst.} \end{cases}.$$

3.2.1 Statische Probitmodelle mit autokorrelierten Störgrößen

Ein großer Vorteil multivariater Probitmodelle liegt, wie bereits angeführt darin, daß sie (im Gegensatz z.B. zum multivariaten Logitmodell) die Modellierung komplexerer Störgrößenprozesse zulassen (die multivariate logistische Verteilung läßt per Annahme nur Korrelationskoeffizienten mit dem Wert 0.5 zu; vgl. Johnson and Kotz, 1972, S. 293f.).

Das vorherrschende Konzept des Varianzkomponentenmodells, welches unterstellt, daß sich die Störgröße u_{it} aus einem individuenspezifischen Effekt α_i und einer über die Zeit

stochastisch unabhängig und identisch verteilten Komponente ε_{it} zusammensetzt (die zeitspezifischen Effekte λ_t werden, falls vorhanden, in einem entsprechend erweiterten Strukturparametervektor $\boldsymbol{\beta}$ berücksichtigt), kann somit bei Verwendung von Probitmodellen verallgemeinert werden. In Analogie zu den linearen ökonometrischen Modellen kann dabei z.B. eine, mit wachsendem zeitlichen Abstand abnehmende, Korrelation der latenten Zufallsvariablen y_{it}^* unterstellt werden.

Gegeben sei das Schwellenwertmodell:

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst} \end{cases},$$

mit $y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}$, $i = 1,\ldots,N;\ t = 1,\ldots,T$, mit $u_{it} = \alpha_i + \varepsilon_{it}$ und mit $y_{it}^*, \boldsymbol{x}_{it}', \boldsymbol{\beta}$ und u_{it} wie oben (vgl. 2.2.2.1) definiert und ε_{it} stochastisch unabhängig und identisch verteilt.

Abweichend davon werde nun für die Störgrößen ε_{it}, mit einem über die Zeit konstanten Mittelwert und einer über die Zeit konstanten Varianz, folgender Markov-Prozeß erster Ordnung unterstellt (vgl. auch Hamerle und Ronning, 1994, S. 423f.) :

$$u_{it} = \alpha_i + \varepsilon_{it} \text{ und } \varepsilon_{\text{it}} = \rho\varepsilon_{i,t-1} + v_{it}\ , \quad i = 1,\ldots,N;\ t = 1,\ldots,T,$$

wobei $|\rho| < 1$, $E(\alpha_i) = 0$, $E(v_{it}) = 0$, $E(\alpha_i^2) = \sigma_\alpha^2$, $E(v_{it}^2) = \sigma_v^2$, $E(v_{it}v_{js}) = 0$ für $i \neq j$ oder $t \neq s$ und $E(\alpha_i\varepsilon_{jt}) = 0$ für alle $i, j = 1,\ldots,N;\ t = 1,\ldots,T$.

Die Identifizierbarkeit wird erreicht mittels der Normierungsvorschrift $Var(u_{it}) = Var(\alpha_i + \varepsilon_{it}) = \sigma_\alpha^2 + \sigma_\varepsilon^2 = 1$. Somit ist (vgl. Greene, 1990, S. 417f.):

$$\sigma_\alpha^2 = 1 - \sigma_\varepsilon^2, \text{ mit } \sigma_\varepsilon^2 = \frac{\sigma_v^2}{1-\rho^2}$$

und für die Kovarianz von u_{it} und u_{is} ergibt sich (vgl. Greene, 1990, S.418):

$$Cov(u_{it}, u_{is}) = E((\alpha_i + \varepsilon_{it})(\alpha_i + \varepsilon_{is})) = E((\alpha_i + \rho\varepsilon_{i,t-1} + v_{it})(\alpha_i + \varepsilon_{is}))$$

$$=\ldots= E((\alpha_i + v_{it} + \rho v_{i,t-1} + \ldots + \rho^{|t-s+1|}v_{i,s+1} + \rho^{|t-s|}\varepsilon_{is})(\alpha_i + \varepsilon_{is}))$$

$$= \sigma_\alpha^2 + \sigma_\varepsilon^2\rho^{|t-s|} = \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{|t-s|}.$$

Damit erhält man die Kovarianzmatrix (wegen $\sigma_u^2 = 1$ ist diese identisch mit der Korrelationsmatrix) des Störgrößenvektors $\boldsymbol{u_i}$ zu:

$$Cov(\boldsymbol{u_i}) = E(\boldsymbol{u_i}\boldsymbol{u_i'}) = \begin{pmatrix} 1 & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & 1 & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} & \cdots & 1 \end{pmatrix}.$$

Die Abhängigkeit der Entscheidungen der Individuen über die Zeit impliziert somit eine T-variate Normalverteilung, und die Likelihoodfunktion dieses Modells beinhaltet folglich für jede Beobachtungseinheit T-fache Normalverteilungsintegrale. Diese sind bereits für kleinere Werte von T schwierig auszuwerten. Es gilt deshalb, Ansätze zu finden, welche die Berechnung T-dimensionaler Integrale bei der ML-Schätzung vermeiden. Mittels dieser Ansätze gilt es, die Parameter $\boldsymbol{\beta}$, σ_α^2 und ρ im Probitmodell mit autokorrelierten Störgrößen konsistent zu schätzen.

Einen möglichen Lösungsansatz für das obige Modell (und multivariate Probitmodelle mit beliebiger, positiv definiter Kovarianzmatrix $\boldsymbol{\Sigma}$) liefert der sogenannte MECOSA-Ansatz (″MEan and COvariance Structure Analysis″; Schepers und Arminger, 1992), auf den in Abschnitt 4.1 eingegangen wird.

3.2.2 Dynamische Probitmodelle mit äquikorrelierten Störgrößen

Man kann, wie einleitend zu Abschnitt 3.2 erwähnt, das Probitmodell mit zufälligen Effekten dadurch verallgemeinern, daß verzögert endogene Variablen als erklärende Variablen zugelassen werden. Dabei wird, wie bereits angeführt, in der hier vorliegenden Arbeit vereinfachend unterstellt, daß nur um eine Periode verzögerte endogene Variable einen signifikanten Einfluß auf y_{it}^* besitzen. Außerdem wird angenommen, daß die Störgrößen, u_{it}, unabhängig von dem Vektor der exogenen Variablen, $\boldsymbol{x_{it}}$, verteilt sind und einem Varianzkomponentenmodell genügen, d.h., $u_{it} = \alpha_i + \varepsilon_{it}$ für alle i,t.

Damit basiert das dynamische Probitmodell auf folgendem Schwellenwertmodell:

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst,} \end{cases}$$

mit $y_{it}^* = \boldsymbol{x_{it}'}\boldsymbol{\beta} + \gamma\, y_{i,t-1} + u_{it}$, mit $u_{it} = \alpha_i + \varepsilon_{it}$, $i = 1,\ldots,N;\, t = 1,\ldots,T$ und mit $\boldsymbol{x_{it}'}$, $\boldsymbol{\beta}$ und u_{it} wie in Abschnitt 2.2.2.1 definiert. γ ist der Parameter der um eine Periode verzögerten endogenen Variablen, $y_{i,t-1}$.

Wenn aber die Wahrscheinlichkeit, daß y_{it} den Wert eins annimmt, von $y_{i,t-1}$ abhängt, so sind zwei wichtige Punkte zu beachten:

Der erste liegt in der Frage, wie man bei einem solchen Markov-Prozeß erster Ordnung mit dem Anfangszustand ′umgeht′; denn die Spezifikation des Anfangszustandes ist in den hier zu betrachtenden dynamischen Probitmodellen entscheidend für die Stichprobeneigenschaften der Parameterschätzwerte.

Hierzu ist festzustellen, daß in vielen empirischen Untersuchungen in den Wirtschafts- und Sozialwissenschaften angenommen wird, daß es sich bei den Anfangszuständen um modellexogene Konstanten handelt (vgl. Hsiao, 1986, S. 169). Diese Annahme ist jedoch nur dann sinnvoll, wenn die Störgrößen des Anfangszustandes und die der Folgeperioden, bzw. die Störgrößen im zugehörigen Schwellenwertmodell, nicht autokorreliert sind oder - im Falle der Autokorrelation - bei Beobachtungsbeginn des Panels zufällig ein neuer Prozeß beginnt, was beides nur selten zutreffen wird. Sind diese Annahmen aber nicht erfüllt, und geht man trotzdem von exogenen Anfangszuständen aus, so kann man mittels der ML-Methode inkonsistente Parameterschätzwerte erhalten (vgl. Hsiao, 1986, S. 170). Allgemein ist es also angezeigt, eine Wahrscheinlichkeitsverteilung für y_{i0} bei gegebenem α_i zu spezifizieren.

Der zweite Punkt betrifft das schon angesprochene Problem der Unterscheidung von Zustandsabhängigkeit und unbeobachteter Heterogenität. Da die u_{it}, trotz der Annahme stochastisch unabhängiger und identisch verteilter Störgrößen ε_{it}, wegen $u_{it} = \alpha_i + \varepsilon_{it}$ eine positive Autokorrelation (σ_α^2) aufweisen, besteht auch eine positive Korrelation zwischen $y_{i,t-1}$ und u_{it}.
Diese Tatsache beruht auf folgender Überlegung: Je größer der Wert der Störgröße $u_{i,t-1}$, desto größer ist die latente Zufallsvariable $y^*_{i,t-1}$, und desto größer ist die Wahrscheinlichkeit, daß $y^*_{i,t-1}$ den Schwellenwert (null) übersteigt und desto größer ist damit die Wahrscheinlichkeit, daß $y_{i,t-1}$ den Wert eins annimmt.

Die Parameterschätzung für einen ausgewählten Zeitpunkt im obigen Modell (mittels eines herkömmlichen Probitansatzes) würde folglich nach oben verzerrte Schätzwerte für den Parameter γ liefern, ein Effekt, der aus der Behandlung von Fehlern-in-den-Variablen-Modellen bekannt ist (vgl. z.B. Schneeweiß, 1971, S. 218ff.). Man bedenke hierzu, daß ein γ signifikant von null verschieden auf eine Zustandsabhängigkeit hinweist, daß aber γ durchaus auch dann als signifikant von null verschieden ermittelt werden kann, wenn in einem statischen Modell eine Autokorrelation der Störgrößen besteht, diese Korrelationsstruktur der Störterme aber nicht angemessen im Modell berücksichtigt wird.

Die Abhängigkeit der Schätzung von dem jeweiligen Anfangszustand läßt sich wie folgt deutlich machen: Bezeichnet man die Dichtefunktion von y_{i0} bei gegebenem α_i mit $f(y_{i0}|\alpha_i)$, so ergibt sich, bei expliziter Berücksichtigung der unbeobachteten individuenspezifischen Effekte in diesem Modell, folgende marginale logarithmierte Likelihoodfunktion (vgl. Hamerle und Ronning, 1994, S. 430):

$$ln\,L = \sum_{i=1}^{N} ln \int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x'_{it}\beta} + \gamma\, y_{i,t-1} + \sigma_\alpha \tilde{\alpha}_i\big)\Big)\, f(y_{i0}|\alpha_i)\, \varphi(\tilde{\alpha}_i)\, d\tilde{\alpha}_i\,.$$

Diese stellt eine Verallgemeinerung der in Abschnitt 2.2.2.1 dargestellten unbedingten logarithmierten Likelihoodfunktion dar, indem nunmehr die Zustandsabhängigkeit ($\gamma\, y_{i,t-1}$) und die Approximation des Anfangszustandes Berücksichtigung finden.

Hierbei werden, wie dort, die individuenspezifischen Effekte als Zufallsvariablen, welche einer Normalverteilung mit Erwartungswert 0 und Varianz σ_α^2 genügen, modelliert, und für $\tilde{\alpha}_i$ gilt:

$$\tilde{\alpha}_i = \frac{\alpha_i}{\sigma_\alpha}\,.$$

Schließlich wird unterstellt, daß y_{i0} durch den gleichen stochastischen Prozeß, der auch die Paneldaten erzeugt hat, determiniert wurde und die Dichtefunktion von y_{i0} bei gegebenem α_i, bezeichnet als $f(y_{i0}|\alpha_i)$.

In praxi ist die Wahrscheinlichkeitsverteilung $f(y_{i0}|\alpha_i)$ allerdings nur schwer herzuleiten, und erhöht somit die Komplexität der Berechnung der logarithmierten Likelihoodfunktion.

Unterstellt man die Anfangszustände y_{i0} dagegen als exogen, so erhält man für das dynamische Probitmodell mit äquikorrelierten Störgrößen (und exogenem Anfangszustand) folgende marginale logarithmierte Likelihoodfunktion:

$$ln\,L = \sum_{i=1}^{N} ln \int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x'_{it}\beta} + \gamma\, y_{i,t-1} + \sigma_\alpha \tilde{\alpha}_i\big)\Big)\, \varphi(\tilde{\alpha}_i)\, d\tilde{\alpha}_i\,.$$

Die Kovarianzmatrix des Störgrößenvektors $\boldsymbol{u_i} = (\mathrm{u}_{i1},\ldots,u_{iN})',\, i = 1,\ldots,N$ besitzt für das dynamische Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand folgende Form:

$$Cov(\boldsymbol{u_i}) = \begin{pmatrix} \sigma_\alpha^2 + \sigma_\varepsilon^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_\alpha^2 & \sigma_\alpha^2 + \sigma_\varepsilon^2 & \cdots & \sigma_\alpha^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 + \sigma_\varepsilon^2 \end{pmatrix},$$

wobei unterstellt wird, daß $\sigma_\alpha^2 + \sigma_\varepsilon^2 = 1$.

In diesem Modell gilt es somit, die Parameter $\boldsymbol{\beta}, \gamma$ und σ_α^2 konsistent zu schätzen.

Als Alternative zur Annahme eines exogenen Anfangszustandes schlägt Heckman (1981a) folgende Approximation des Anfangszustandes y_{i0} vor:

$$y_{i0}^* = \boldsymbol{x}_{i0}'\boldsymbol{\delta} + u_{i0} \text{ und } y_{i0} = \begin{cases} 1 & \text{falls } y_{i0}^* > 0 \\ 0 & \text{sonst} \end{cases}, \quad i = 1,\ldots,N \ .$$

Man unterstellt dabei, daß die Störgröße u_{i0} mit α_i und ε_{it} beliebig korreliert sein kann, und $E(u_{i0}\,u_{it}) = \sigma_{0t}$ für alle t sei. Erweitert man den Vektor der latenten Variablen $\boldsymbol{y}_i^*$ um y_{i0}^*, so erhält man, unter den Annahmen bzgl. der Störgrößen und ohne weitere Restriktionen, folgende Kovarianzmarix der Ordnung $T+1$ für den Störgrößenvektor $\tilde{\boldsymbol{u}}_i = (u_{i0}, u_{i1}, \ldots, u_{iT})'$ (vgl. Hamerle und Ronning, 1984, S. 431):

$$Cov(\tilde{\boldsymbol{u}}_i) = E(\tilde{\boldsymbol{u}}_i\tilde{\boldsymbol{u}}_i') = \begin{pmatrix} \sigma_{u0}^2 & \sigma_{01} & \sigma_{02} & \cdots & \sigma_{0T} \\ \sigma_{01} & \sigma_\alpha^2 + \sigma_\varepsilon^2 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_{02} & \sigma_\alpha^2 & \ddots & \sigma_\alpha^2 & \vdots \\ \vdots & \vdots & \sigma_\alpha^2 & \ddots & \sigma_\alpha^2 \\ \sigma_{0T} & \sigma_\alpha^2 & \cdots & \cdots & \sigma_\alpha^2 + \sigma_\varepsilon^2 \end{pmatrix}.$$

Um die Parameter in diesem Modell zu identifizieren, bedarf es zusätzlicher Restriktionen, beispielsweise der Form $\boldsymbol{\sigma}_{u0}^2 = 1$ und $\boldsymbol{\sigma}_u^2 = \boldsymbol{\sigma}_{\boldsymbol{\alpha}}^2 + \boldsymbol{\sigma}_{\boldsymbol{\varepsilon}}^2 = 1$. Damit lautet die Kovarianzmatrix (Korrelationsmatrix) der $\tilde{\boldsymbol{u}}_i$ (vgl. Hamerle und Ronning, 1994, S. 431):

$$Cov(\tilde{\boldsymbol{u}}_i) = E(\tilde{\boldsymbol{u}}_i\tilde{\boldsymbol{u}}_i') = \begin{pmatrix} 1 & \sigma_{01} & \sigma_{02} & \cdots & \sigma_{0T} \\ \sigma_{01} & 1 & \sigma_\alpha^2 & \cdots & \sigma_\alpha^2 \\ \sigma_{02} & \sigma_\alpha^2 & \ddots & \sigma_\alpha^2 & \vdots \\ \vdots & \vdots & \sigma_\alpha^2 & \ddots & \sigma_\alpha^2 \\ \sigma_{0T} & \sigma_\alpha^2 & \cdots & \cdots & 1 \end{pmatrix}.$$

Im dynamischen Probitmodell mit äquikorrelierten Störgrößen und approximiertem Anfangszustand gilt es somit, die Parameter $\boldsymbol{\beta}, \gamma, \delta, \sigma_\alpha^2$ und $\sigma_{01}, \ldots, \sigma_{0T}$ konsistent zu schätzen.

Bezüglich möglicher Auswirkungen einer Vernachlässigung der hier bestehenden Heterogenität der $\boldsymbol{u}_i$, bzw. der $\tilde{\boldsymbol{u}}_i$ zeigt Heckman mittels eines Beschäftigungsmodells für Frauen ′mittleren Alters′, daß eine Nichtberücksichtigung dieser Heterogenität zu eklatanten Verzerrungen der Parameterschätzwerte führen kann. Seine Studie behandelte u.a. den Effekt der Arbeitserfahrung auf eine zukünftige Beschäftigung. Ließ Heckman im vorgegebenen Fall die bestehende Heterogenität außer acht, so wurde der Effekt der Arbeitserfahrung um das bis zu Zehnfache überschätzt (zitiert in: Hsiao, 1986, S. 175ff.).

In dem nun folgenden Abschnitt werden die Modellvarianten aus den Abschnitten 3.2.1 und 3.2.2 miteinander verknüpft; man gelangt damit zu einem dynamischen Probitmodell mit autokorrelierten Störgrößen.

3.2.3 Dynamische Probitmodelle mit autokorrelierten Störgrößen

Dynamische Probitmodelle mit autokorrelierten Störgrößen stellen eine Erweiterung der in Abschnitt 3.2.2 betrachteten dynamischen Probitmodelle mit äquikorrelierten Störgrößen dar.

Ausgehend von der auch den Ausführungen des Abschnitts 3.2.2 zugrundeliegenden Modellspezifikation

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst} \end{cases} \quad , \quad i = 1,\dots,N;\, t = 1,\dots,T \ ,$$

und $y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \gamma\, y_{i,t-1} + u_{it}$, mit $u_{it} = \alpha_i + \varepsilon_{it}$ und mit $\boldsymbol{x}_{it}', \boldsymbol{\beta}, \gamma, \alpha_i$ und ε_{it} wie dort definiert, wird für ε_{it} nunmehr ein Markov-Prozeß erster Ordnung unterstellt:

$$\varepsilon_{\mathrm{it}} = \rho\varepsilon_{i,t-1} + v_{it} \ , \quad i = 1,\dots,N;\, t = 1,\dots,T \, .$$

Zugleich gelten die in Abschnitt 3.2.1 getroffenen Annahmen bezüglich des Zusammenhangs der Störkomponenten $\alpha_i, \varepsilon_{it}$ und v_{it}. Damit ergibt sich die dort spezifizierte Kovarianzmatrix von $\boldsymbol{u_i}$ zu (vgl. Abschnitt 3.2.1):

$$Cov(\boldsymbol{u_i}) = E(\boldsymbol{u_i}\boldsymbol{u_i}') = \begin{pmatrix} 1 & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & 1 & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} & \cdots & 1 \end{pmatrix},$$

wobei man aus Identifikationsgründen unterstellte, daß $\sigma_u^2 = \sigma_\alpha^2 + \sigma_\varepsilon^2 = 1$.

In diesem Modell sind somit die Parameter $\boldsymbol{\beta}$, γ, σ_α^2 und ρ konsistent zu schätzen.

Das Ziel bei der empirischen Analyse eines dynamischen Probitmodells mit autokorrelierten Störgrößen ist es, auf Zustandsabhängigkeit und gleichzeitig auf Autokorrelation der Störgrößen zu testen. Zum Test auf Zustandsabhängigkeit benötigt man dabei einen unverzerrten und konsistenten Schätzer für den Koeffizienten γ der Variablen $y_{i,t-1}$.

Bei der empirischen Analyse des dynamischen Probitmodells mit äquikorrelierten Störgrößen erhält man einen unverzerrten und konsistenten Schätzer für den Parameter γ, indem man die marginale logarithmierte Likelihoodfunktion, die nicht mehr von den individuenspezifischen Effekten $\boldsymbol{\alpha}_i$ abhängt, nach dem Parameter γ maximiert (vgl. Abschnitt 3.2.2).

Das Vorgehen bei der Bildung der marginalen logarithmierten Likelihoodfunktion, d.h., die Multiplikation der sich ergebenden logarithmierten Likelihoodfunktion mit der Dichtefunktion für α_i und die anschließende Integration bezüglich α_i, führt im dynamischen Probitmodell mit autokorrelierten Störgrößen nicht in gleicher Weise zum gewünschten Erfolg, da hier die Störgrößen ε_{it} zusätzlich einem Markov-Prozeß (erster Ordnung) unterliegen und damit autokorreliert sind.

Die Variable $y_{i,t-1}$ weist somit eine positive Korrelation mit der Störgröße u_{it} auf (vgl. Abschnitt 3.2.2); eine Schätzung des Parameters γ mit Hilfe der marginalen logarithmierten Likelihoodfunktion würde nach oben verzerrte Schätzwerte liefern. Im dynamischen Probitmodell mit autokorrelierten Störgrößen ist es demnach nicht möglich, mittels der Maximierung der marginalen logarithmierten Likelihoodfunktion auf Zustandsabhängigkeit und gleichzeitig auf Autokorrelation der Störgrößen zu testen.

Zusammenfassend hat sich in diesem Abschnitt gezeigt, daß die ′wünschenswerte′ Modellspezifikation eines dynamischen Probitmodells mit autokorrelierten Störgrößen, mit den in der vorliegenden Arbeit betrachteten Verfahren, nicht konsistent schätzbar ist.

Bevor es im Kapitel 5 um die Datengrundlage der empirischen Analyse statischer und dynamischer Panelmodelle geht, werden im Kapitel 4 einige Grundlagen der Implementierung dieser Modelle dargelegt.

4 Grundlagen der Implementierung statischer und dynamischer Probitmodelle

Der Abschnitt 2.3 befaßte sich mit verschiedenen Aspekten der ML-Schätzung in binomialen Antwortmodellen, sowie mit allgemeinen Grundlagen der Implementierung statischer und dynamischer Panelmodelle. Diese Ausführungen werden nun vertieft bzw. erweitert.

Im ersten Abschnitt des vorliegenden Kapitels wird ein spezifischer Lösungsansatz für die ML-Schätzung der im Abschnitt 3.2.1 dargestellten Modellvariante, einem statischen Probitmodell mit autokorrelierten Störgrößen, diskutiert: der MECOSA-Ansatz (Abschnitt 4.1). Dies ist notwendig, da ein Probitmodell mit autokorrelierten Störgrößen einer Schätzung mit den ′üblichen′ Maximierungsalgorithmen nicht zugänglich ist.

Abschnitt 4.2 befaßt sich mit Lösungsalgorithmen für (dynamische) Probitmodelle mit äquikorrelierten Störgrößen (vgl. Abschnitt 2.2.2.1 und Abschnitt 3.2.2). In diesen Modellen ist die zu maximierende logarithmierte Likelihoodfunktion im allgemeinen nicht global konkav; mittels der Quasi-Newton-Methoden (vgl. Fletcher, 1980), der Davidon-Fletcher-Powell (DFP)- sowie der Broyden-Fletcher-Goldfarb-Shanno (BFGS)-Methode, Weiterentwicklungen der Newton-Raphson-Methode (vgl. Abschnitt 2.3.1), gelingt jedoch auch eine Maximierung nicht global konkaver Funktionen. Für diesen Fall besitzen sie gegenüber der Newton-Raphson-Methode verbesserte Konvergenzeigenschaften (vgl. Greene, 1997, S. 203f.).

Das zusätzliche Problem, daß die oben genannte logarithmierte Likelihoodfunktion pro Beobachtungseinheit ein Integral über das Produkt von T (Standard-) Normalverteilungsfunktionen enthält, kann mit Hilfe der Gauß-Hermite-Quadratur (vgl. Bronstein, Semendjajev, 1987, S. 763f.) numerisch effizient gelöst werden. Dieses Verfahren wird in Abschnitt 4.3 behandelt.

4.1 Schätzung statischer Probitmodelle mit autokorrelierten Störgrößen - Der MECOSA-Ansatz

Für das statische Probitmodell mit autokorrelierten Störgrößen (vgl. Abschnitt 3.2.1) beinhaltet die zugehörige logarithmierte Likelihoodfunktion T-dimensionale Normalverteilungsintegrale, welche bereits für kleinere Werte von T schwierig zu berechnen sind. Einen einfachen Lösungsweg bietet der MECOSA-Ansatz, der nur die Lösung bivariater Normalverteilungsintegrale erforderlich macht.

Im Abschnitt 4.1.1 wird zunächst das im Abschnitt 3.2.1 spezifizierte Probitmodell in die Nomenklatur des MECOSA-Ansatzes überführt; im Abschnitt 4.1.2 wird die Strategie der Parameterschätzung mittels des MECOSA-Ansatzes dargestellt.

4.1.1 Modellspezifikation des statischen Probitmodells mit autokorrelierten Störgrößen in der Nomenklatur des MECOSA-Ansatzes

Das dem statischen Probitmodell mit autokorrelierten Störgrößen zugrundeliegende Schwellenwertmodell wurde bisher in folgender Weise formuliert:

$$y_{it} = \begin{cases} 1 & \text{falls } y_{it}^* > 0 \\ 0 & \text{sonst} \end{cases},$$

mit $y_{it}^* = \boldsymbol{x}_{it}'\boldsymbol{\beta} + u_{it}$, $i = 1,\ldots,N;\ t = 1,\ldots,T$ und mit $\boldsymbol{x}_{it}', \boldsymbol{\beta}$ und u_{it} wie oben definiert.

In dieser Modellspezifikation wurde der Störgrößenvektor $\boldsymbol{u_i} = (u_{i1},\ldots,u_{iT})'$ als T-variat normalverteilt mit Erwartungswertvektor $\boldsymbol{0}$ und beliebiger Kovarianzmatrix $\boldsymbol{\Sigma}$ unterstellt.

Bei dem hier betrachteten Probitmodell mit autokorrelierten Störgrößen definierte man $\boldsymbol{\Sigma}$ wie folgt (vgl. Abschnitt 3.2.1):

$$Cov(\boldsymbol{u_i}) = \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho & 1 & \cdots & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-1} & \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{T-2} & \cdots & 1 \end{pmatrix}$$

Eine Zusammenfassung über alle T Wellen führt zu folgender Schreibweise:

$$\boldsymbol{y_i^*} = \boldsymbol{\Pi} \cdot \boldsymbol{x_i} + \boldsymbol{u_i}\ , \quad i = 1,\ldots,N\ ,$$

mit $\boldsymbol{\Pi}$, der Koeffizientenmatrix der Ordnung $T \times (T \cdot K)$ und definiert als das Kronecker-Produkt $\boldsymbol{\Pi} = \boldsymbol{I}_{\mathrm{T}} \otimes \boldsymbol{\beta}'$ und dem Vektor der Regressoren $\boldsymbol{x_i} = (\boldsymbol{x_{i1}'},\ldots,\boldsymbol{x_{iT}'})'$ mit Dimension $(T \cdot K)$.

Im Zusammenhang mit der Parameterschätzung ist in diesem Modell demnach eine logarithmierte Likelihoodfunktion zu maximieren, welche pro Beobachtungseinheit (pro Individuum) ein T-dimensionales Normalverteilungsintegral enthält. Dieses ist wie folgt aufgebaut (vgl. Greene, 1997, S. 897):

$$ln\, L = \sum_{i=1}^{N} ln\, P(\boldsymbol{y_i})\text{, mit } P(\boldsymbol{y_i}) = \int_{C_1}\!\!\ldots \int_{C_t}\!\!\ldots \int_{C_T} \varphi(z_1,\ldots,z_t,\ldots,z_T)\, dz_T \ldots dz_t \ldots dz_1\,,$$

mit $\boldsymbol{\varphi}(z_1,\ldots,z_T)$ der Dichtefunktion der T-variaten Standardnormalverteilung. Die Integrationsbereiche C_t sind dabei folgendermaßen definiert:

$$C_t = \begin{cases} \{c_t \text{ mit } -\infty < c_t \leq \boldsymbol{x'_{it}\beta}\} & \text{falls } y_{it} = 1 \\ \{c_t \text{ mit } \boldsymbol{x'_{it}\beta} < c_t\} & \text{falls } y_{it} = 0 \end{cases}.$$

Diese Integrationsbereiche ergeben sich unter der Annahme des Schwellenwertes null und angesichts der Symmetrie der Normalverteilung aufgrund folgender Überlegungen: Es ist

$$P(y_{it} = 1) = P(y_{it}^* > 0) = P(\boldsymbol{x'_{it}\beta} + u_{it} > 0) = P(u_{it} > -\boldsymbol{x'_{it}\beta}) = P(u_{it} < \boldsymbol{x'_{it}\beta})$$

und

$$P(y_{it} = 0) = \ldots = P(u_{it} > \boldsymbol{x'_{it}\beta}).$$

Der MECOSA-Ansatz umgeht nun die direkte Lösung des T-dimensionalen Normalverteilungsintegrals durch Verwendung eines dreistufigen Schätzverfahrens, welches konsistente Parameterschätzer liefert. Diese Schätzstrategie wird im folgenden Abschnitt dargestellt.

4.1.2 Parameterschätzung mittels des MECOSA-Algorithmus

Die Schätzstrategie des MECOSA-Ansatzes ist eine modifizierte ML-Methode, basierend auf der multivariaten Normalverteilung der $\boldsymbol{y_i^*}$. Küsters (1987) zeigt, daß das dreistufige Verfahren dieses Lösungsansatzes konsistente und asymptotisch normalverteilte Schätzer liefert.

Dabei wird, ausgehend vom obigen Schwellenwertmodell, in der ersten Stufe des Schätzverfahrens für jede einzelne Welle ein statisches Probitmodell geschätzt. Aus Identifikationsgründen geht man dabei davon aus, daß die Varianz der u_{it} in den verschiedenen Wellen gleich eins ist ($\sigma_t^2 = 1$ für alle $t = 1,\ldots,T$; vgl. Abschnitt 2.2). Damit erhält man konsistente Schätzwerte von $\boldsymbol{\beta}_t$ für alle $t = 1,\ldots,T$.

Die zweite Stufe des MECOSA-Schätzverfahrens besteht aus zwei Teilschritten. Im ersten Teilschritt werden die $\frac{T(T-1)}{2}$ unbekannten Elemente v_{ts}, $s,t = 1,\ldots,T$, der Kovarianzmatrix $\boldsymbol{\Sigma}$ geschätzt. Dieses geschieht zunächst, ohne den Kovarianzen (v_{ts}) Restriktionen aufzuerlegen (abgesehen von der für die Identifizierbarkeit der jeweiligen Strukturparametervektoren $\boldsymbol{\beta}_t, t = 1,\ldots,T$ notwendigen Normierung der Varianzen).

Da die bivariaten Normalverteilungen von y_{it}^* und y_{is}^* ($s,t = 1,\ldots,T$ und $s \neq t$) vollständig durch die Mittelwerte und die Varianzen von y_{it}^* und y_{is}^*, sowie durch deren Kovarianzen

festgelegt sind, erhält man die (unrestringierten) ML-Schätzer $\hat{\nu}_{ts}$ für die $\frac{T(T-1)}{2}$ unbekannten Kovarianzen dadurch, daß man die jeweiligen bivariaten logarithmierten Likelihoodfunktionen von y_{it} und y_{is} bzgl. ν_{ts} maximiert.

Im einzelnen ergibt sich für diesen ersten Teilschritt der zweiten Stufe des MECOSA-Verfahrens folgendes:

Durch Einsetzen der Parameterschätzwerte $\hat{\boldsymbol{\beta}}_t$ und $\hat{\boldsymbol{\beta}}_s$ für die Wellen t und s (aus der ersten Stufe) in die gemeinsame logarithmierte Likelihoodfunktion der binären Zufallsvariablen y_{it} und y_{is} gelangt man zu der bivariaten logarithmierten Likelihoodfunktion

$$ln\, L(\nu_{ts}) = \sum_{i=1}^{N} ln\, P(y_{it}, y_{is} \,|\, \boldsymbol{x_i}, \hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s)\,.$$

Die zugehörige bivariate Normalverteilungsfunktion für y_{it} und y_{is} hat dabei (beispielhaft) folgendes Aussehen:

$$P(y_{it} = 1, y_{is} = 1 \,|\, \boldsymbol{x_i}, \hat{\boldsymbol{\beta}}_t, \hat{\boldsymbol{\beta}}_s) = \int_{-\infty}^{\boldsymbol{x}_{it}'\hat{\boldsymbol{\beta}}_t} \int_{-\infty}^{\boldsymbol{x}_{is}'\hat{\boldsymbol{\beta}}_s} \varphi(z_t, z_s)\, dz_s\, dz_t\,,$$

mit $\boldsymbol{\varphi}(z_t, z_s)$, der Dichtefunktion der bivariaten Standardnormalverteilung:

$$\boldsymbol{\varphi}(z_t, z_s) = \frac{1}{2\pi \cdot \sqrt{1-\nu_{ts}^{\;2}}} \cdot exp\left(-\frac{1}{2\cdot(1-\nu_{ts}^{\;2})} \left(z_t^2 + z_s^2 - 2\cdot \nu_{ts} \cdot z_t \cdot z_s \right) \right).$$

Die Maximierung der bivariaten logarithmierten Likelihoodfunktion bzgl. ν_{ts} liefert in diesem ersten Schritt der zweiten Stufe konsistente Parameterschätzwerte, $\hat{\nu}_{ts}$, für die Elemente ν_{ts} der Kovarianzmatrix $\boldsymbol{\Sigma}$.

Im zweiten Teilschritt der zweiten Stufe werden die in der ersten Stufe des Verfahrens für die T Wellen ermittelten Schätzwerte $\hat{\boldsymbol{\beta}}_t,\, t = 1,\ldots,T$ und die $\frac{T(T-1)}{2}$ im ersten Teilschritt der zweiten Stufe ermittelten Kovarianzschätzwerte $\hat{\nu}_{ts}$ in einem $\left[(T\cdot K) + \frac{T(T-1)}{2} \right]$-dimensionalen Vektor $\hat{\boldsymbol{\kappa}}$ zusammengefaßt. Schließlich wird in diesem Teilschritt ein konsistenter Schätzer $\hat{W}$ für die asymptotische Kovarianzmatrix W von $\hat{\boldsymbol{\kappa}}$ ermittelt. Letzteres geschieht mittels analytischer Ableitungen erster Ordnung der T logarithmierten Likelihoodfunktionen der ersten Stufe, sowie numerischer Ableitungen zweiter Ordnung der $\frac{T(T-1)}{2}$ bivariaten logarithmierten Likelihoodfunktionen des ersten Teilschrittes der zweiten Stufe (vgl. Küsters, 1987, S. 55f.).

Die fundamentale Annahme des MECOSA-Ansatzes wird in der dritten Stufe deutlich. Diese Annahme besteht darin, daß die Einflüsse der jeweiligen erklärenden Variablen als über alle T Wellen konstant angesehen werden, m.a.W., es wird von einer über alle Wellen konstanten Parameterstruktur ausgegangen. Dies geschieht durch die Einführung entsprechender Parameterrestriktionen in der dritten Stufe des Algorithmus. Zusätzlich wird im zugehörigen Schwellenwertmodell eine, mit dem Faktor ρ abnehmende, Korrelation der Störgrößen unterstellt. Auch diese Annahme (vgl. Abschnitt 3.2.1) wird in der dritten Stufe der MECOSA-Schätzung durch Einführung entsprechender Restriktionen gewährleistet. Es gilt also:

$$\boldsymbol{\beta}_1 = \ldots = \boldsymbol{\beta}_T = \boldsymbol{\beta} \text{ und } v_{ts} = \sigma_\alpha^2 + (1-\sigma_\alpha^2)\rho^{|t-s|}, \text{ für } t \neq s \text{ und } t,s > 0.$$

Faßt man die zu schätzenden Parameter in einem Vektor $\boldsymbol{\vartheta}$, mit $\boldsymbol{\vartheta} = (\boldsymbol{\beta}', \sigma_\alpha^2, \rho)$ zusammen, so kann die Minimierungsvorschrift wie folgt geschrieben werden:

$$Q(\boldsymbol{\vartheta}) = (\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\vartheta}))' \, \hat{W}^{-1} \, (\hat{\boldsymbol{\kappa}} - \boldsymbol{\kappa}(\boldsymbol{\vartheta})) \, .$$

Die Schätzung der Parameter erfolgt damit mittels eines ′Minimum distance′-Ansatzes (vgl. Arminger, 1994, S.8), wobei man den Wert von $\boldsymbol{\vartheta}$ sucht, der, gewichtet mit der Matrix $\hat{\boldsymbol{W}}^{-1}$, den Abstand von $\hat{\boldsymbol{\kappa}}$ und $\boldsymbol{\kappa}(\boldsymbol{\vartheta})$ minimiert. Auf diese Weise erhält man konsistente Schätzwerte für $\boldsymbol{\beta}$, σ_α^2 und ρ.

4.2 Quasi-Newton-Methoden

Wie bereits angeführt, ist die zu maximierende Likelihoodfunktion bei (dynamischen) Probitmodellen mit äquikorrelierten Störgrößen nicht notwendig global konkav. Die Davidon-Fletcher-Powell (DFP)- sowie die Broyden-Fletcher-Goldfarb-Shanno (BFGS)-Methode sind zwei Verfahren der nichtlinearen Optimierung zur (iterativen) Bestimmung der ML-Schätzer in Probitmodellen mit nicht notwendigerweise global konkaver logarithmierter Likelihoodfunktion. Beide Verfahren stellen Weiterentwicklungen der Newton-Raphson-Methode dar (vgl. Abschnitt 2.3.1). Die nachfolgenden Ausführungen zeigen die Unterschiede dieser Ansätze auf.

Für eine allgemein, global konkave Funktion $F(\boldsymbol{\theta})$ liefert die Newton-Raphson-Methode in den meisten Fällen den bestmöglichen Algorithmus, da dieser Algorithmus unabhängig vom Startwert für den Parametervektor $\boldsymbol{\theta}$ gegen den Wert $\hat{\boldsymbol{\theta}}$ konvergiert, der die Funktion $F(\boldsymbol{\theta})$ maximiert. Für diese Optimierung der Funktion $F(\boldsymbol{\theta})$ werden dazu die partiellen ersten und zweiten Ableitungen der Funktion $F(\boldsymbol{\theta})$ nach $\boldsymbol{\theta}$ benötigt. Die Newton-Raphson-Methode beruht dabei auf folgendem Algorithmus (vgl. Abschnitt 2.3.1):

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_t - \boldsymbol{H}^{-1}(\boldsymbol{\theta}_k) \, \boldsymbol{g}(\boldsymbol{\theta}_k) \, ,$$

mit $\boldsymbol{g}(\boldsymbol{\theta}_k)$, dem Gradienten von $F(\boldsymbol{\theta}_k)$, $\boldsymbol{H}(\boldsymbol{\theta}_k)$ der Matrix der partiellen zweiten Ableitungen der Funktion $F(\boldsymbol{\theta}_k)$ nach $\boldsymbol{\theta}_k$ und k, dem Iterationsindex.

Im Falle einer global konkaven Funktion $F(\boldsymbol{\theta})$ ist die zugehörige Hesse-Matrix $\boldsymbol{H}$ negativ definit, das (-1)-fache der Inversen der Hesse-Matrix - die sogenannte Informationsmatrix - ist somit positiv definit.
Eine nicht global konkave Funktion $F(\boldsymbol{\theta})$ verlangt jedoch in der Regel alternative Iterationsmethoden, da weit vom Optimum entfernte Punkte ($\boldsymbol{\theta}_k$) dazu führen können, daß die Hesse-Matrix ($\boldsymbol{H}(\boldsymbol{\theta}_k)$) nicht negativ definit ist. Dieser Tatbestand kann zur Folge haben, daß die Newton-Raphson-Methode nicht konvergiert (vgl. Greene, 1997, S. 203).

Daher verwendet man im Falle nicht notwendig global konkaver Zielfunktionen sogenannte Quasi-Newton-Methoden. Diese ähneln dem Algorithmus der Newton-Raphson-Methode, unterscheiden sich aber von diesem dadurch, daß nicht das (-1)-fache der Inversen der Hesse-Matrix ($-\boldsymbol{H}^{-1}(\boldsymbol{\theta}_{k+1})$), sondern eine Matrix $\boldsymbol{W}_{k+1}$ zur Bestimmung von $\boldsymbol{\theta}_{k+2}$ verwendet wird:

$$\boldsymbol{\theta}_{k+2} = \boldsymbol{\theta}_{k+1} + \boldsymbol{W}_{k+1}\boldsymbol{g}(\boldsymbol{\theta}_{k+1}).$$

Sowohl die DFP- als auch die BFGS-Methode benötigen zur Herleitung dieser Matrix $\boldsymbol{W}_{k+1}$ und damit zur nichtlinearen Optimierung der Funktion $F(\boldsymbol{\theta})$, ausschließlich die partiellen ersten Ableitungen der Funktion $F(\boldsymbol{\theta})$. Beide Algorithmen zeigen auch für nicht global konkave Funktionen $F(\boldsymbol{\theta})$ ausgezeichnete Konvergenzeigenschaften, da die Matrix $\boldsymbol{W}_{k+1}$, im Gegensatz zum (-1)-fachen der Inversen der Hesse-Matrix, in jedem Iterationsschritt positiv definit ist (vgl. Greene, 1997, S. 204).

Die DFP-Methode verwendet folgenden Algorithmus zur Bestimmung der Matrix $\boldsymbol{W}_{k+1}$ (vgl. Greene, 1997, S. 204):

$$\boldsymbol{W}_{k+1} = \boldsymbol{W}_k + \frac{\boldsymbol{\delta}_k\boldsymbol{\delta}'_k}{\boldsymbol{\delta}'_k\boldsymbol{\gamma}_k} + \frac{\boldsymbol{W}_k\boldsymbol{\gamma}_k\boldsymbol{\gamma}'_k\boldsymbol{W}_k}{\boldsymbol{\gamma}'_k\boldsymbol{W}_k\boldsymbol{\gamma}_k},$$

mit $\boldsymbol{\delta}_k = \boldsymbol{W}_k \cdot \boldsymbol{g}(\boldsymbol{\theta}_k)$ und $\boldsymbol{\gamma}_k = \boldsymbol{g}(\boldsymbol{\theta}_{k+1}) - \boldsymbol{g}(\boldsymbol{\theta}_k)$.

Dabei bezeichnet $\boldsymbol{g}(\boldsymbol{\theta}_k)$ den Gradienten der zu optimierenden Funktion $F(\boldsymbol{\theta}_k)$ und k den Iterationsindex. Für die ′Startmatrix′ $\boldsymbol{W}_0$ wird im allgemeinen eine Einheitsmatrix gleicher Ordnung verwendet.

Indem die beiden obigen Quotienten in jeder Iterationsstufe zu $\boldsymbol{W}_k$ addiert werden, ergibt sich die (in jeder Iterationsstufe) positiv definite Matrix $\boldsymbol{W}_{k+1}$. Nach einer entsprechenden Anzahl von Iterationsschritten ist $\boldsymbol{W}_{k+1}$ approximativ gleich dem (-1)-fachen der Inversen der Hesse-Matrix ($-\boldsymbol{H}^{-1}$) (vgl. Greene, 1997, S.204).

Die BFGS-Methode stellt eine Weiterentwicklung der DFP-Methode dar, und ist in vielen Fällen effizienter als die DFP-Methode; die zugehörige Matrix $\boldsymbol{W_{t+1}}$ berechnet sich in diesem Fall wie folgt (vgl. Greene, 1997, S. 204f.):

$$\boldsymbol{W}_{k+1} = \boldsymbol{W}_k + \frac{\boldsymbol{\delta}_k \boldsymbol{\delta}'_k}{\boldsymbol{\delta}'_k \boldsymbol{\gamma}_k} + \frac{\boldsymbol{W}_k \boldsymbol{\gamma}_k \boldsymbol{\gamma}'_k \boldsymbol{W}_k}{\boldsymbol{\gamma}'_k \boldsymbol{W}_k \boldsymbol{\gamma}_k} - v_k \, \boldsymbol{d}_k \boldsymbol{d}'_k ,$$

wobei $v_k = \boldsymbol{\gamma}_k{}' \boldsymbol{W}_k \boldsymbol{\gamma}_k$ und $\boldsymbol{d}_k = \left(\frac{1}{\boldsymbol{\delta}'_k \boldsymbol{\gamma}_k} \right) \boldsymbol{\delta}_k - \left(\frac{1}{\boldsymbol{\gamma}_k{}' \boldsymbol{W}_k \boldsymbol{\gamma}_k} \right) \boldsymbol{W}_k \, \boldsymbol{\gamma}_k$.

Auch hier wird als Startmatrix im allgemeinen eine Einheitsmatrix verwendet. Der Algorithmus zur Bestimmung der Matrix $\boldsymbol{W_{k+1}}$ nach der BFGS-Methode unterscheidet sich deshalb vom Algorithmus der DFP-Methode nur durch den Term $v_k \, \boldsymbol{d}_k \boldsymbol{d}'_k$.

4.3 Die Gauss-Hermite-Quadratur

Die ML-Schätzung in statischen und dynamischen Probitmodellen mit äquikorrelierten Störgrößen erfordert die Lösung von Integralen der folgenden Form:

$$\int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1) \big(\boldsymbol{x}'_{it} \boldsymbol{\beta} + \sigma_\alpha \tilde{\alpha}_i \big) \Big) \, \boldsymbol{\varphi}(\tilde{\alpha}_i) \, d\tilde{\alpha}_i$$

für das statische Probitmodell (vgl. Abschnitt 2.2.2.1), beziehungsweise

$$\int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1) \big(\boldsymbol{x}'_{it} \boldsymbol{\beta} + \gamma \, y_{i,t-1} + \sigma_\alpha \tilde{\alpha}_i \big) \Big) \, f(y_{i0} | \alpha_i) \, \boldsymbol{\varphi}(\tilde{\alpha}_i) \, d\tilde{\alpha}_i$$

für das dynamische Probitmodell (vgl. Abschnitt 3.2.2).
Diese Integrale können mit Hilfe der Gauß-Hermite-Quadratur numerisch effizient berechnet werden.

Unter (numerischer) Quadratur versteht man dabei die ″näherungsweise zahlenmäßige Berechnung des Wertes eines eindimensionalen bestimmten Integrals $\int_a^b f(x) dx$ aus der Kenntnis einzelner Funktionswerte des Integranden und gewisser allgemeiner Eigenschaften desselben.″ (Bronstein und Semendjajev, 1987, S. 761).

Insbesondere dient die Gauß-Hermite-Quadratur der numerisch effizienten Berechnung von Integralen der Form

$$\int_{-\infty}^{\infty} exp(-z^2) \cdot g(z) \, dz .$$

Die approximative Lösung derartiger Integrale erfolgt auf der Grundlage folgenden Ausdrucks (vgl. Butler and Moffitt, 1982, S. 762):

$$\int_{-\infty}^{\infty} exp(-z^2) \cdot g(z)\, dz = \sum_{j=1}^{G} w_j \cdot g(z_j)\,,$$

mit z_j, der j-ten Stützstelle, w_j, dem zugehörigen Gewicht und G, der Gesamtanzahl der Stützstellen (vgl. Bronstein, Semendjajev, 1987, S. 761ff.).

Damit läßt sich auch ein Integral der Form

$$\int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \sigma_\alpha \tilde{\alpha}_i\big)\Big)\, \boldsymbol{\varphi}(\tilde{\alpha}_i)\, d\tilde{\alpha}_i$$

mit Hilfe der Gauß-Hermite-Quadratur numerisch effizient lösen, da die Dichtefunktion der Standardnormalverteilung $\boldsymbol{\varphi}$ einen Term der Form $\exp(-z^2)$ enthält. In diesem Fall ergibt sich folgendes Integral:

$$\int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \sigma_\alpha \tilde{\alpha}_i\big)\Big) \cdot \frac{1}{\sqrt{2\pi}} exp(-\tfrac{1}{2} \cdot \tilde{\alpha}_i^2)\, d\tilde{\alpha}_i$$

und durch Substitution von $h_i = \frac{1}{\sqrt{2}} \cdot \tilde{\alpha}_i$ erhält man schließlich:

$$\int_{-\infty}^{\infty} \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \sqrt{2} \cdot \sigma_\alpha \cdot h_i\big)\Big) \cdot \frac{1}{\sqrt{2\pi}} exp(-h_i^2)\, \sqrt{2} \cdot d\, h_i\,.$$

Auf dieses Integral kann die oben angegebene Approximation angewendet werden, wobei sich die Funktion $g(z)$ aus dem folgenden Produkt von T Standardnormalverteilungsfunktionen zusammensetzt:

$$g(z) = \frac{1}{\sqrt{\boldsymbol{\pi}}} \cdot \prod_{t=1}^{T} \boldsymbol{\Phi}\Big((2 \cdot y_{it} - 1)\big(\boldsymbol{x}_{it}'\boldsymbol{\beta} + \sqrt{2} \cdot \sigma_\alpha \cdot z\big)\Big)\,.$$

Für die numerisch effiziente Berechnung des jeweiligen Integrals benötigt man die G Stützstellen z_j und die zugehörigen Gewichte w_j; diese Angaben findet man in Abramowitz and Stegun (1964, S. 924).

5 Datengrundlage der empirischen Analyse des Arbeitsstatus männlicher Erwerbspersonen

Die Daten für die quantitative Analyse mittels der in den vorangegangenen Abschnitten behandelten statischen und dynamischen nichtlinearen Panelmodelle stammen aus dem Datensatz für die Wellen B bis N (1985-1997) des Deutschen Sozio-ökonomischen Panels (SOEP; für eine detaillierte Beschreibung siehe Wagner, Schupp und Rendtel, 1991). Die Zielpopulation dieses Panels ist die wohnberechtigte Bevölkerung in der Bundesrepublik Deutschland einschließlich Berlin-West (Gebietsstand vor dem 03. 10. 1990).

Mittels dieser Daten, soll eine empirische Analyse der individuellen Arbeitslosigkeit männlicher Arbeiter und Angestellter unter Verwendung von Probitmodellen vorgenommen werden. Es gelangen deshalb nur die Männer in die Stichprobe, die im Basisjahr 1984 zwischen 25 und 51 Jahren alt waren, und für die die benötigten Daten für alle dreizehn Wellen vorliegen, Beamte, Selbständige, Auszubildende und Rentner wurden definitionsgemäß ausgeschlossen. Damit umfaßt das Panel 527 Personen. Da insbesondere Rentner und Auszubildende nicht in die Stichprobe aufgenommen werden, erscheint es weitgehend unproblematisch, alle ′Nicht-Beschäftigungs-Zustände′ der Probanden als ′Arbeitslosigkeit′ zu bezeichnen.

Das SOEP beinhaltet neben Informationen über den Beschäftigungsstatus beobachteter Individuen zum jeweiligen Interviewzeitpunkt, auch Informationen über die Anzahl der Monate, während der das betrachtete Individuum in den jeweiligen Vorperioden als arbeitslos gemeldet war. Weiterhin liefert das SOEP Informationen über die Anzahl der ′Arbeitslosigkeitszustände′ der befragten Personen zwischen 1974 und 1984 und die Gesamtdauer dieser Zustände (in Monaten) für diesen Zeitraum.

Daneben stellt das SOEP Informationen über sozio-ökonomische Charakteristika der betrachteten Bevölkerungsgruppe zur Verfügung. Hierzu zählen u. a. bestimmte Haushaltscharakteristika, wie das Nettohaushaltseinkommen des betrachteten Haushalts, der Familienstand der betrachteten Person und die Anzahl der Kinder, für die der Haushalt im jeweiligen Vorjahr Kindergeld bezog. Außerdem werden vom SOEP persönliche Charakteristika der Befragten, wie Nationalität, Alter, Gesundheitszustand, Ausbildung und Nettoverdienst erfaßt und bereitgestellt.

Im folgenden werden die in den empirisch zu analysierenden Probitmodellen berücksichtigten Variablen zunächst definiert; anschließend wird darauf eingegangen, weshalb die einzelnen Variablen ausgewählt wurden und welcher Einfluß (Richtung) der jeweiligen Variablen auf den Arbeitsstatus der betrachteten Personen zu erwarten ist.

Abschließend wird auf Besonderheiten des Datensatzes eingegangen, und es werden zu erwartende Probleme bei der empirischen Analyse der nichtlinearen Panelmodelle mittels des hier beschriebenen Datensatzes angesprochen.

5.1 Definition der Variablen und deren erwarteter Einfluß auf den Arbeitsstatus

Die abhängige Variable ist in allen zu betrachtenden Probitmodellen der Arbeitsstatus der betrachteten Person:

EST Die endogene Variable EST bezeichnet den Arbeitssstatus des betrachteten Individuums zum jeweiligen Zeitpunkt des Interviews; hierbei handelt es sich um eine binäre Zufallsvariable, welche den Wert eins annimmt, falls das Individuum zum Zeitpunkt des Interviews arbeitslos ist, andernfalls ist EST = 0.

Im einzelnen werden in die, im nächsten Kapitel empirisch zu analysierenden Probitmodelle die folgenden exogenen Variablen aufgenommen:

ALT Die Variable ALT mißt das Lebensalter der betrachteten Person (in Jahren).

ALTSQ Die Variable ALTSQ definiert als $ALTSQ := ALT^2 / 100$ wird verwendet, um einen erwarteten nichtlinearen Zusammenhang von Arbeitsstatus und Lebensalter zu berücksichtigen (Lebenszyklus).

GZ Die Variable GZ gibt an, ob die betrachtete Person schwerbehindert ist (GZ = 1) oder nicht (GZ = 0).

BB1 Die Variable BB1 nimmt den Wert eins an, falls die betrachtete Person eine abgeschlossene Berufsausbildung besitzt; andernfalls ist BB1 = 0.

BB2 Die Variable BB2 nimmt den Wert eins an, falls die betrachtete Person einen Universitätsabschluß vorweisen kann; andernfalls ist BB2 = 0.

BST2 Die Variable BST2 nimmt den Wert eins an, falls die betrachtete Person als Angestellter geführt wird; andernfalls (die Person wird als Arbeiter geführt) ist BST2 = 0.

NAT Die Variable NAT kennzeichnet die Nationalität der betrachteten Person. Dabei wird nur zwischen Ausländern (NAT = 1) und Deutschen (NAT = 0) unterschieden.

FST Die Variable FST kennzeichnet den Familienstand der betrachteten Person. Ist der Arbeitnehmer verheiratet, dann ist FST = 1; andernfalls ist FST = 0.

CHILD Die Variable CHILD quantifiziert die Anzahl der Kinder, für die der Haushalt der befragten Person im Vorjahr Kindergeld bezogen hat.

HINC Die Variable HINC bezeichnet die Differenz aus dem Nettoeinkommen aller Haushaltsmitglieder des jeweiligen Haushaltes (inklusive Bezüge, also inklusive Wohngeld, Kindergeld, BAföG, etc.) und dem Nettoarbeitseinkommmen des Haushaltsvorstandes (in 1000 DM).

ALQ Die Variable ALQ kennzeichnet die Arbeitslosigkeitsquote des Bundeslandes, in dem die betrachtete Person im jeweiligen Jahr ihren Wohnsitz hatte (in Prozent).

ALD24 Die Variable ALD24 steht für die Dauer der Arbeitslosigkeit des beobachteten Merkmalsträgers innerhalb der jeweils zurückliegenden zwei Jahre (in Monaten).

ESTL Die Variable ESTL bezeichnet den Arbeitssstatus des Merkmalsträgers in der Vorperiode. Sie nimmt den Wert eins an, falls der Merkmalsträger in der Vorperiode arbeitslos war; andernfalls ist ESTL = 0.

Die folgenden drei Variablen sind Indikatoren der Vorgeschichte des Arbeitsstatus der jeweiligen Person in den Jahren 1974 bis 1984 und dienen im dynamischen Probitmodell mit Äquikorrelation (vgl. Abschnitt 6.6) der Approximation des Anfangszustandes:

ALD Die Variable ALD kennzeichnet die Dauer der Arbeitslosigkeit der betrachteten Person im Zeitraum von 1974 bis 1984 (in Monaten).

ALDSQ Die Verwendung von $ALDSQ = ALD^2 / 100$ wird analog der Variablen ALTSQ begründet.

ALH Die Variable ALH gibt Auskunft über die absolute Häufigkeit der Arbeitslosigkeit der betrachteten Person im Zeitraum von 1974 bis 1984.

Als wesentliche Bestimmungsgrößen der individuellen Arbeitslosigkeit bzw. deren Eintreff-Wahrscheinlichkeit werden somit persönliche Charakteristika, wie Alter, Gesundheitszustand, Nationalität, Ausbildung, zurückliegende Zeiten der Arbeitslosigkeit der betrachteten Person, daneben Haushaltscharakteristika, wie Familienstand, Nettohaushaltseinkommen aller Haus-

haltsmitglieder und die Anzahl der Kinder, für die der Haushalt im Vorjahr Kindergeld bezogen hat, gesehen. Ergänzt wird die Liste der Bestimmungsgründe durch eine Angabe zur Arbeitsmarktlage in dem jeweiligen Bundesland.

Diese Variablen wurden ausgewählt, da sie entweder die Entscheidung der jeweiligen Person, Arbeit anzubieten, beeinflussen und/oder Effekte der ′Nachfrageseite′ individueller Beschäftigung wiederspiegeln. Mehrere Variablen stehen sowohl für Effekte der ′Angebotsseite′, als auch für Effekte der ′Nachfrageseite′ individueller Beschäftigung.

So kann unterstellt werden, daß die Höhe des jeweiligen ′Nettohaushaltseinkommens′ der betrachteten Person nur deren Wahrscheinlichkeit, Arbeit anzubieten, beeinflußt. Andererseits spielen die Variablen ′Alter′, ′Gesundheitszustand′, ′Nationalität′ und ′Familienstand′ sowohl beim Arbeitsangebot, als auch bei der Arbeitsnachfrage eine Rolle, während die Ausbildung einer Person und die Tatsache, ob es sich um einen Arbeiter oder Angestellten handelt, eher für Effekte der ′Nachfrageseite′ individueller Beschäftigung von Bedeutung sind.

Hinsichtlich der Einflüsse einzelner Variablen auf den Arbeitsstatus ist folgendes denkbar: Es wird ein nichtlinearer Einfluß des (Lebens-) Alters auf die Wahrscheinlichkeit, arbeitslos zu sein, erwartet. Deshalb wird außer der Variablen ALT auch die Variable ALTSQ in den zu betrachtenden Probitmodellen berücksichtigt. Dabei wird davon ausgegangen, daß die Beschäftigungswahrscheinlichkeit zunächst mit zunehmendem Alter und damit zunehmender Berufserfahrung steigt, dieser Effekt sich jedoch für ältere Arbeitnehmer umkehren kann. Für den Koeffizienten des linearen Terms erwartet man deshalb ein negatives, für den Koeffizienten des quadratischen Terms ein positives Vorzeichen.

Weiterhin ist davon auszugehen, daß Arbeitnehmer mit einer abgeschlossenen Berufsausbildung eine geringere Disposition zeigen, arbeitslos zu werden, als solche ohne eine abgeschlossene Berufsausbildung. Der Koeffizientenschätzwert für die hier definierte Variable ′abgeschlossene Berufsausbildung′ (BB1) sollte also negativ sein. Dieser Zusammenhang wird auch für die Variable ′Universitätsabschluß′ (BB2) unterstellt.

Verheiratete Männer zeigen, so wird hier unterstellt, ein geringeres Risiko, arbeitslos zu werden, als ledige Männer; denn einerseits kann angenommen werden, daß mit einer Heirat und der damit einhergehenden gestiegenen Verantwortung der Anreiz zu arbeiten steigt, andererseits ist zu vermuten, daß ledige Männer in ′Krisensituationen′ eher entlassen werden. Für den Koeffizienten dieser Variablen (FST) wird also ein negatives Vorzeichen erwartet.

Vorzeichen und Ausmaß des Einflusses der exogenen Variablen ′Schwerbehinderung′ (GZ) und ′Nationalität′ (NAT) lassen sich schwer vorhersagen, da sich für beide Variablen sowohl Gründe für einen positiven Effekt auf die Beschäftigungswahrscheinlichkeit, als auch Gründe für einen negativen Effekt auf diese Wahrscheinlichkeit anführen lassen.

Eine geringere Wahrscheinlichkeit, arbeitslos zu werden, wird hingegen für Personen erwartet, die im Angestelltenverhältnis beschäftigt sind (im Vergleich zu Arbeitern). Somit ist von einem negativen Koeffizientenschätzwert für die Variable ′Angestellter/Arbeiter′ (BST2) auszugehen.

Ein wichtiger Indikator für die individuelle Entscheidung, Arbeit anzubieten, dürfte auch das ′Nettohaushaltseinkommen′ (HINC) sein. Das Vorliegen eines hohen Nettohaushaltseinkommens (Differenz aus dem Nettoeinkommen aller Haushaltsmitglieder des jeweiligen Haushaltes, einschließlich der Bezüge und dem Nettoarbeitseinkommmen des Haushaltsvorstandes) läßt einen negativen Einfluß auf die Arbeitswilligkeit/-notwendigkeit des jeweiligen Individuums erwarten, da mit steigendem Nettohaushaltseinkommen der Anreiz, Arbeit anzubieten, sinken kann. So könnte eine als ausreichend angesehene Summe aus Bezügen (z.B. Arbeitslosengeld und BAföG-Zahlungen) und ein Nettoarbeitsverdienst der Ehefrau Gründe für die beobachtete Person sein, keine Arbeitsleistung anzubieten.

Effekte der ′Nachfrageseite′ werden auch durch Einbeziehung der Variablen ′Arbeitslosenquote′ (ALQ) im jeweiligen Bundesland (und im jeweiligen Jahr) berücksichtigt. Dies wird damit begründet, daß mit steigender Arbeitslosenquote eine größere Wahrscheinlichkeit, arbeitslos zu sein, zu erwarten ist, da die Arbeitsnachfrage offenbar gesunken ist.

Der Effekt der zur Dynamisierung des Modells führenden Berücksichtigung der verzögert endogenen Variablen (ESTL) dürfte positiv sein, was zugleich auf eine Zustandsabhängigkeit hinweisen würde. Als gleichgerichtet wird auch der Effekt der Variablen ALD24, der Dauer der Arbeitslosigkeit während des jeweils zurückliegenden Zeitraums (6 bis 30 Monate vor dem jeweiligen Interview), eingestuft.

5.2 Besonderheiten des Datensatzes und daraus zu erwartende Schätzprobleme

Der hier verwendete Datensatz weist folgende Besonderheiten auf:

- Die Arbeitslosenquote in der durch das Panel erfaßten Stichprobe ist in allen betrachteten Wellen deutlich geringer als die jeweilige Arbeitslosenquote in der Grundgesamtheit. Eine Erklärung für diese Tatsache liegt möglicherweise darin begründet, daß Arbeitslose grundsätzlich weniger motiviert sind, über einen längeren Zeitraum an einer Panelbefragung teilzunehmen.

- Die Panelstichprobe beinhaltet nur wenige Individuen (weniger als fünf Prozent), die einen Universitätsabschluß besitzen. Ein möglicher Grund für diese Unterrepräsentanz liegt in der ausschließlichen Berücksichtigung von Arbeitern und Angestellten.
- Bei der Auswertung der Angaben für die Variable ′Nettohaushaltseinkommen′ (HINC) ergaben sich teilweise negative Werte. Dies ist deshalb bemerkenswert, weil diese Variable definiert ist als die Differenz aus dem Nettohaushaltseinkommen aller Haushaltsmitglieder einschließlich der Bezüge und dem Nettoarbeitseinkommen der betrachteten Person. Da das Nettoarbeitseinkommen des Haushaltsvorstandes per definitionem ein Teil des Nettohaushaltseinkommens ist, sollte die Variable HINC nur positive Werte annehmen. Die negativen Werte für die Variable HINC lassen sich deshalb nur als falsche Angabe der befragten Person im jeweiligen Personen- bzw. Haushaltsfragebogen des SOEP interpretieren, wobei die Gründe für diese falschen Angaben hier nicht bekannt sind. Das diesbezügliche Vorgehen bei der Erstellung des Datensatzes war deshalb folgendes: Überstieg das Nettoarbeitseinkommen des Haushaltsvorstandes das Nettoeinkommen des Haushaltes in einer der dreizehn Wellen des Panels um mehr als zehn Prozent, so wurde diese Person in der Panelstichprobe nicht berücksichtigt. War die Differenz negativ, aber dem Betrag nach geringer als zehn Prozent des Nettohaushaltseinkommens, so wurde die Variable HINC der betreffenden Person in der jeweiligen Welle gleich null gesetzt.

Bei der empirischen Analyse einzelner Wellen sind außerdem folgende Schätzprobleme zu erwarten:

- In Wellen mit geringer Arbeitslosenquote in der Panelstichprobe sind die zu erwartenden Schätzergebnisse mit Vorsicht zu interpretieren; denn, je weniger Personen in einer bestimmten Welle arbeitslos sind, desto weniger Informationen für die empirische Analyse der individuellen Arbeitslosigkeit enthält der Datensatz.
- Da der Stichprobenumfang des einzelnen Querschnitts (der einzelnen Welle) mit 527 Personen relativ gering ist, kann hier der Fall eintreten, daß bestimmte Kombinationen von Werten binärer exogener Variablen mit den Werten der endogenen Variable nicht besetzt sind. Beispielsweise kann es passieren, daß keine Person mit einem Universitätsabschluß arbeitslos ist. In diesem Fall ist es nicht möglich, einen Schätzwert für den Koeffizienten der jeweiligen Variablen zu bestimmen.

6 Ergebnisse der Schätzung

Die empirische Analyse beschränkt sich auf die Wellen B bis N (1985 bis 1997), da das Deutsche Sozio-ökonomische Panel die notwendige Information zur Erzeugung der Variablen ′Arbeitslosigkeitsdauer′ (ALD24) für die Welle A nicht bereitstellt.

6.1 Ein statisches Probitmodell für die einzelnen Wellen

Die quantitative Analyse, getrennt nach Wellen, wird hier als eine erste Annäherung an die bei der quantitativen Analyse der übrigen statischen und dynamischen Probitmodelle zu erwartenden Ergebnisse gesehen. Die Parameterschätzwerte für dieses Modell sind in den Tabellen 6.1.1 und 6.1.2 zusammengefaßt:

- Ausgehend von der Tatsache, daß das Alter der untersuchten Personen zu Beginn des Panels zwischen 25 und 51 Jahren lag, zeigt sich in allen Wellen, in denen der Einfluß dieser Variablen signifikant nachweisbar ist, daß die Wahrscheinlichkeit, arbeitslos zu sein, mit zunehmendem Alter der Arbeiter und Angestellten zunächst sinkt und anschließend wieder zunimmt.[1] Bemerkenswert ist auch, daß ein Einfluß in den Jahren 1992 und 1993 signifikant nicht nachweisbar ist. Dies ist möglicherweise darauf zurückzuführen, daß die Arbeitslosenquote in der Panelstichprobe in diesen Jahren, mit Werten von 3,80 und 4,36 Prozent, vergleichsweise gering ist.

- Eine Schwerbehinderung des Arbeitnehmers (GZ = 1) erhöht dessen Wahrscheinlichkeit, arbeitslos zu sein, nur in den Jahren 1988 und 1989 signifikant. Im Jahr 1995 ist dagegen ein signifikant positiver Effekt auf die Wahrscheinlichkeit, beschäftigt zu sein, nachweisbar. In den übrigen (zehn) Jahren ist ein signifikanter Einfluß dieser Variablen nicht nachweisbar.

- Ein deutlicher Einfluß der Variablen ′abgeschlossene Berufsausbildung′ (BB1) dahingehend, die Chancen arbeitslos zu sein zu reduzieren, zeigt sich nur in den Jahren 1986 und 1987, während die abgeschlossene Berufsausbildung im Jahr 1997 eher ′kontraproduktiv′ wirkte. Möglicherweise ist dieses Ergebnis auf die Tatsache zurückzuführen, daß in der betrachteten Panelstichprobe nur etwa sieben Prozent der Arbeiter und Angestellten keine abgeschlossene Berufsausbildung aufweisen.

---

[1] Der Graph der Funktion, die diesen Zusammenhang beschreibt, hat also die Form einer nach oben geöffneten Parabel. Für die in den folgenden Abschnitten betrachteten statischen und dynamischen Probitmodelle wird jeweils angegeben, ab welchem Alter die Wahrscheinlichkeit wieder zunimmt, arbeitslos zu werden.

Tabelle 6.1.1:

Statisches Probitmodell

Ergebnisse der Parameterschätzung für die einzelnen Wellen

Wellen B (1985) bis H (1991)

	Welle B		Welle C		Welle D		Welle E		Welle F		Welle G		Welle H	
Variable	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert
ALT	-0.1801	-4.441	-0.1265	-3.526	-0.1329	-3.334	-0.1832	-3.917	-0.1700	-3.822	-0.1396	-2.919	-0.0785	-2.173
ALTSQ	0.2415	4.045	0.1649	3.044	0.1716	2.851	0.1807	2.862	0.2199	3.436	0.2009	2.778	0.0741	1.348
GZ	0.8029	1.933	0.8683	1.742	0.6755	1.436	1.4316	3.206	1.2218	3.234	0.6517	1.173	0.0675	0.147
BB1	0.1931	0.428	-0.7239	-2.030	-0.8013	-2.092	0.0167	0.039	0.2294	0.499	-0.2582	-0.431	-0.3321	-0.689
BST2	-1.0408	-2.632	-1.1518	-2.272	-1.2190	-2.283	-1.5169	-2.934	-0.8189	-2.321	-1.1207	-1.635	-0.3145	-0.802
NAT	0.0793	0.210	-0.2554	-0.646	0.3368	0.727	0.1267	0.223	-1.6908	-2.502	-0.6739	-1.139	-0.0252	-0.052
FST	-0.0303	-0.088	-0.3165	-0.859	0.2647	0.598	0.8006	1.542	-0.4445	-1.063	-0.7974	-1.431	-0.7720	-1.717
CHILD	-0.0750	-0.879	0.0894	0.933	0.0106	0.106	0.1460	1.357	-0.0286	-0.214	0.0750	0.490	0.1167	1.139
HINC	0.4568	4.478	0.2711	3.181	0.3717	3.764	0.4441	4.282	0.2758	2.656	0.2955	2.561	0.1501	1.675
ALQ	0.1055	2.152	0.1113	2.112	0.0547	1.020	0.1136	1.952	0.1425	1.981	0.0133	0.163	0.0374	0.542
ALD24	0.0779	3.598	0.1260	5.017	0.1535	5.679	0.1456	5.206	0.0756	2.777	0.2069	5.687	0.1500	5.746
LOG-L	-81.024		-71.156		-58.634		-57.278		-51.760		-53.564		-54.487	

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Es ist LOG-L der Wert der logarithmierten Likelihoodfunktion ausgewertet beim Maximum der Funktion.

Quelle: Eigene Berechnungen

Tabelle 6.1.2:

Statisches Probitmodell

Ergebnisse der Parameterschätzung für die einzelnen Wellen

Wellen I (1992) bis N (1997)

	Welle I		Welle J		Welle K		Welle L		Welle M		Welle N	
Variable	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert	Parameter-Schätzwert	Asympt. t-Wert
ALT	-0.0114	-0.233	-0.0455	-0.841	-0.1027	-2.279	-0.1241	-2.376	-0.0984	-2.235	-0.1633	-3.222
ALTSQ	-0.0243	-0.338	0.0648	0.883	0.1469	2.477	0.1643	2.522	0.0946	1.690	0.1844	3.108
GZ	0.1333	0.288	0.0880	0.169	-0.1135	-0.275	-1.5313	-1.954	0.5526	1.783	0.3741	1.078
BB1	0.2031	0.273	-0.2250	-0.414	-0.2769	-0.579	0.9577	1.223	0.7739	1.279	1.5364	1.861
BST2	-0.6635	-1.205	-0.7354	-1.238	-0.9858	-1.768	-1.2857	-1.575	-1.2194	-2.501	-1.8434	-2.822
NAT	0.1120	0.184	-1.1421	-2.157	-0.7354	-1.329	-0.6628	-1.271	-0.0771	-0.289	-0.0749	-0.249
FST	-0.3452	-0.605	-1.4590	-2.929	-0.6212	-1.236	-1.0758	-2.042	-0.2063	-0.490	-1.1280	-2.912
CHILD	-0.0693	-0.510	0.0565	0.364	0.1621	1.359	-0.2062	-1.216	0.0238	0.215	-0.1567	-1.156
HINC	0.1120	1.020	0.0995	1.037	0.2885	3.844	0.2742	3.344	0.2985	4.592	0.4247	4.536
ALQ	-0.2245	-2.015	-0.0618	-0.557	-0.0415	-0.496	-0.0469	-0.548	-0.0570	-0.886	0.0221	0.340
ALD24	0.2022	5.297	0.1949	5.565	0.1854	5.776	0.2121	6.783	0.1509	7.949	0.1560	8.268
LOG-L	-57.782		-52.255		-50.377		-57.312		-66.365		-54.462	

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Es ist LOG-L der Wert der logarithmierten Likelihoodfunktion ausgewertet beim Maximum der Funktion.

Quelle: Eigene Berechnungen

- Auf die Berücksichtigung der Variablen ′Universitätsabschluß′ (BB2) wurde in den hier betrachteten Modellen verzichtet, da in der Mehrzahl der Wellen alle betrachteten Personen, die einen Universitätsabschluß vorweisen konnten, beschäftigt waren. Diese Tatsache machte eine Schätzung des zugehörigen Koeffizienten dieser Variablen in den oben genannten Wellen unmöglich (vgl. dazu auch Abschnitt 5.2). In den übrigen Wellen zeigte ein vorhandener Universitätsabschluß einer Person keinen signifikanten Einfluß auf deren Wahrscheinlichkeit, beschäftigt zu sein.

- Der Einfluß der Form des Beschäftigungsverhältnisses ′Angestellter/Arbeiter′ (BST2) auf die Wahrscheinlichkeit, arbeitslos zu sein, ist in allen Wellen negativ; wenngleich während der Jahre 1990 bis 1995 weniger stark ausgeprägt. Dieses Ergebnis zeigt, daß Angestellte in der Mehrzahl der Fälle eine signifikant größere Beschäftigungschance besitzen als Arbeiter.

- Der Einfluß der nationalen Herkunft der Arbeitnehmer (NAT) auf den Arbeitsstatus ist nur schwach ausgeprägt; in den Jahren 19989 und 1993 zeigen jedoch Arbeitnehmer fremder Nationalität das signifikant geringere Risiko, arbeitslos zu werden, als deutsche Arbeitnehmer.

- Der Einfluß des Familienstandes (FST) reduzierte nur in einzelnen Jahren (1993, 1995 und 1997) das Risiko, arbeitslos zu werden, d.h., in diesen Jahren besaßen verheiratete Männer ein signifikant geringeres Risiko, arbeitslos zu werden, als ledige Männer.

- Ein signifikanter Einfluß der ′Anzahl der Kinder′ (CHILD) auf die Wahrscheinlichkeit, beschäftigt zu sein, kann in keinem der Fälle nachgewiesen werden.

- Hinsichtlich der Variablen ′Nettohaushaltseinkommen′ (HINC) zeigt sich (ausgenommen vielleicht in den Jahren 1991 bis 1993) ein deutlicher positiver Zusammenhang zwischen der Höhe des Nettohaushaltseinkommens und der Wahrscheinlichkeit, arbeitslos zu sein.

- Der Einfluß des makroökonomischen Umfeldes auf den Arbeitsstatus, repräsentiert durch die Variable ′Arbeitslosenquote′ (ALQ), ist nur in den ersten Jahren (1985, 1986, 1988 und 1989) des Untersuchungszeitraums signifikant nachweisbar. In dieser Zeit erhöhte eine hohe Arbeitslosenquote auch die Chance, arbeitslos zu werden (Der negative Einfluß dieser Variablen im Jahr 1992 ist kaum interpretierbar).

- Demgegenüber sind die Parameterschätzwerte der Variablen ′Arbeitslosigkeitsdauer′ (ALD24) in allen dreizehn Wellen signifikant positiv, d.h., die Dauer der Arbeitslosigkeit in den jeweils zurückliegenden zwei Jahren erhöht in allen dreizehn Wellen deutlich die Wahrscheinlichkeit, auch in der Folgeperiode arbeitslos zu werden. Der Parameter der Variablen ′Arbeitslosigkeitsdauer′ (ALD24) kann jedoch für die einzelnen Wellen nur dann

unverzerrt und konsistent geschätzt werden, wenn man von nicht autokorrelierten Störgrößen im Schwellenwertmodell ausgehen kann. Andernfalls besteht eine positive Korrelation zwischen der Variablen ALD24 und der zugehörigen Störgröße[1], was zu einer Überschätzung des zur Variablen ALD24 gehörigen Parameters führen kann (vgl. S. 34).

Insgesamt sind die Schätzergebnisse der Probitmodelle für die einzelnen Wellen jedoch mit Vorsicht zu interpretieren, da die einzelnen Querschnitte (527 Personen) nicht sehr breit sind und zudem die Arbeitslosenquote in einzelnen Querschnitten gering ist. Die für die einzelnen Wellen erhaltenen Schätzergebnisse liefern jedoch Anhaltspunkte zur Beurteilung der Schätzergebnisse statischer und dynamischer Probitmodelle, die im folgenden zur Parameterschätzung verwendet werden.

6.2 Ein statisches Probitmodell unter der Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit

Auf die Problematik der vergleichsweise geringen Arbeitslosenquote in der Panelstichprobe für die Parameterschätzung wurde bereits hingewiesen. Die Tabelle 6.2.1 enthält Angaben über die Veränderungen des Arbeitsstatus der betrachteten Individuen im Zeitablauf.

Bezogen auf den Umfang der Panelstichprobe (N = 527) schwankte der Anteil der Individuen, der einen veränderten Arbeitsstatus von einer Welle zur nächsten aufweist, zwischen Werten von 2,3 (Welle I zu Welle J) und 5,7 (Welle L zu Welle M) Prozent. Im Durchschnitt über alle Wellen liegt dieser Anteil bei etwa vier Prozent, d.h., nur jede fünfundzwanzigste Person weist einen veränderten Arbeitsstatus von einer Welle zur nächsten auf.

Sowohl die vergleichsweise geringe Arbeitslosenquote im betrachteten Panel, als auch der geringe Anteil der betrachteten Personen, die einen veränderten Arbeitsstatus von einer Welle

---

[1] Diese Tatsache beruht auf folgender Überlegung: Die Variable ′Arbeitslosigkeitsdauer′ (ALD24) in der Periode t hängt positiv vom Arbeitsstatus der Vorperiode $y_{i,t-1}$ ab. Je größer nun der Wert der Störgröße $u_{i,t-1}$, desto größer ist die latente Zufallsvariable $y^*_{i,t-1}$, und desto größer ist die Wahrscheinlichkeit, daß $y^*_{i,t-1}$ den Schwellenwert (null) übersteigt und desto größer ist damit die Wahrscheinlichkeit, daß $y_{i,t-1}$ den Wert eins annimmt. Besteht aber eine positive Korrelation zwischen den Störgrößen $u_{i,t-1}$ und $u_{i,t}$, so weist auch die Variable ALD24 eine positive Korrelation mit der Störgröße $u_{i,t}$ auf. Diese Argumentation läßt sich in analoger Weise für den Arbeitsstatus der Periode $t-2$, $y_{i,t-2}$ führen.

Tabelle 6.2.1:

Anzahl der Veränderungen des Arbeitsstatus der Individuen beim Übergang von einer Welle zur nächsten

Wellen B (1985) bis N (1997)

	Änderungen gesamt	Übergang in die Arbeitslosigkeit	Übergang in die Beschäftigung
Welle B	21	17	4
Welle C	23	10	13
Welle D	20	9	11
Welle E	17	10	7
Welle F	21	6	15
Welle G	19	9	10
Welle H	20	13	7
Welle I	22	9	13
Welle J	12	6	6
Welle K	23	16	7
Welle L	23	15	8
Welle M	30	22	8
Welle N	24	18	6

Tabelle 6.2.2

Statisches Probitmodell unter der Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit

Ergebnisse der Parameterschätzung

Wellen B (1985) bis N (1997)

Bezeichnung der Variablen	Parameter-Schätzwert	Asymptotischer t-Wert
ALT	-0.1099	-12.532
ALTSQ	0.1221	9.944
GZ	0.3688	3.638
BB1	-0.1183	-1.002
BST2	-0.9636	-7.681
NAT	-0.1464	-1.472
FST	-0.2332	-2.507
CHILD	0.0050	0.178
HINC	0.2466	11.947
ALQ	0.0456	3.273
ALD24	0.1369	23.157
LOG-L	-803.999	

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Es ist LOG-L der Wert der logarithmierten Likelihoodfunktion ausgewertet beim Maximum der Funktion.

Quelle: Eigene Berechnungen

zur nächsten aufweisen, unterstreichen die Notwendigkeit des Übergangs von Querschnitts- zu Panelmodellen.

Deshalb wird im zweiten Schritt der empirischen Analyse ein statisches Probitmodell (vgl. Abschnitt 2.2), unterstellt, d.h., die Daten der dreizehn Wellen (Wellen B bis N) werden zu einem binomialen Antwortmodell zusammengefaßt. Ergebnisse der Parameterschätzung für dieses statische Probitmodell, das vollkommene Unabhängigkeit der individuellen Entscheidungen über die Wellen B bis N (1985 bis 1997) unterstellt, sind in Tabelle 6.2.2 zusammengefaßt.

6.2.1 Ergebnisse der Schätzung eines statischen Probitmodells unter der Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit

Für dieses statische Probitmodell über die Wellen B bis N (mit $13 \cdot 527 = 6851$ Beobachtungen) zeigt die überwiegende Zahl der im Modell berücksichtigten Variablen einen signifikanten Einfluß auf die Wahrscheinlichkeit, arbeitslos zu sein (werden).

- Aus den Vorzeichen der signifikanten Schätzwerte für die Koeffizienten der Altersvariablen ALT und ALTSQ kann - wie schon bei der Analyse einzelner Wellen - geschlossen werden, daß das Arbeitslosigkeitsrisiko mit zunehmendem Alter zunächst sinkt, dann aber, ab einer bestimmten Altersgrenze, wieder zunimmt. Durch Substitution der entsprechenden Werte der Altersvariablen ALT und ALTSQ in die Funktion $f(\text{ALT},\text{ALTSQ}) = \hat{\beta}_{\text{ALT}} \cdot \text{ALT} + \hat{\beta}_{\text{ALTSQ}} \cdot \text{ALTSQ}$ erkennt man, daß die Werte der Funktion $f(\text{ALT},\text{ALTSQ})$ bis zu einem Alter der betrachteten Personen von 45 Jahren fallen, danach jedoch wieder ansteigen.

- Der Koeffizientenschätzwert für die Variable ′Schwerbehinderung′ (GZ) ist signifikant positiv, d.h., Schwerbehinderte besitzen gegenüber nicht behinderten Personen ein signifikant größeres Risiko, arbeitslos zu werden. (Für die einzelnen Wellen war ein eindeutiger Nachweis des Einflusses der Variablen ′Schwerbehinderung′ (GZ) auf die Beschäftigungswahrscheinlichkeit nicht zu erkennen.)

- Ein gesicherter Einfluß einer abgeschlossenen Berufsausbildung (BB1) ist auch hier nicht feststellbar.

- Die Form des Arbeitsverhältnisses ′Angestellter/Arbeiter′ (BST2) wirkt sich signifikant auf die Chancen, arbeitslos zu werden, aus, und zwar derart, daß Angestellte gegenüber Arbeitern ein signifikant geringeres Arbeitslosigkeitsrisiko besitzen.

- Angehörige einer fremden Nationalität zeigen ein eher geringeres Risiko, arbeitslos zu werden, als ihre deutschen Konkurrenten um den Arbeitsplatz. Allerdings ist der Einfluß

der Variablen ′Nationalität′ (NAT) auf die Wahrscheinlichkeit, arbeitslos zu sein, nicht signifikant nachweisbar.

- Verheiratete Männer (FST = 1) besitzen gegenüber ledigen Männern eine signifikant geringere Wahrscheinlichkeit, arbeitslos zu werden. (Über die einzelnen Wellen war der Einfluß des Familienstandes weniger deutlich ausgeprägt.)

- Die Anzahl der Kinder (CHILD) zeigt auch im statischen Probitmodell über die Wellen B bis N keinen signifikanten Einfluß auf die Wahrscheinlichkeit, arbeitslos zu sein.

- Der Einfluß des Nettohaushaltseinkommens (HINC) auf die Chance, arbeitslos zu werden, ist erwartungsgemäß positiv und signifikant von null verschieden. Verfügt demnach der Haushalt der betrachteten Person außerhalb seines Arbeitseinkommens über hohe Einkünfte, so sinkt offenbar der Arbeitsanreiz für den Haushaltsvorstand.

- Auch im statischen Probitmodell über die Wellen B bis N zeigt sich, daß eine höhere Arbeitslosenenquote (ALQ) in einem Bundesland, die Beschäftigungschancen einer Person dieses Bundeslandes signifikant senkt, ein Zusammenhang, der bei der Analyse der einzelnen Wellen nur für den Zeitraum 1985 bis 1989 nachgewiesen werden konnte.

- Auch hier ist der positive Einfluß der Arbeitslosigkeitsdauer (ALD24) auf das Risiko, arbeitslos zu werden, signifikant nachweisbar, d.h., je länger eine betrachtete Person innerhalb der zurückliegenden zwei Jahre arbeitslos war, desto größer ist deren Risiko, ′heute′ arbeitslos zu sein bzw. in der Zukunft arbeitslos zu werden. (vgl. dazu auch entsprechende Ausführungen in Abschnitt 6.1).

Zusammenfassend ist damit festzustellen:

Die Parameterschätzergebnisse des hier betrachteten statischen Probitmodells für das Panel über alle Wellen B bis N spiegeln im allgemeinen die Schätzergebnisse der Probitmodelle für die einzelnen Wellen wieder.

Das hier betrachtete Probitmodell über die Wellen B bis N geht von vollkommener Unabhängigkeit der individuellen Entscheidungen über alle Wellen aus. Diese Annahme ist jedoch sehr restriktiv und nicht notwendigerweise gerechtfertigt.

Mit dem im Abschnitt 6.3 verwendeten statischen Probitmodell mit äquikorrelierten Störgrößen wird diese Frage wieder aufgegriffen und es wird getestet, ob die Hypothese vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit verworfen werden muß oder nicht. Zunächst aber wird im Abschnitt 6.2.2 das Ergebnis eines Likelihood-Quotienten-Tests auf Gleichheit der Parametervektoren der dreizehn Wellen B bis N wiedergegeben.

6.2.2 Ergebnisse des Likelihood-Quotienten-Tests auf Gleichheit der Parametervektoren für die Wellen B bis N

Das statische Probitmodell über alle Wellen (vgl. Abschnitt 6.2.1) dient, neben den Probitmodellen für die einzelnen Wellen (vgl. Abschnitt 6.1), als Grundlage eines Likelihood-Quotienten (LR)-Tests auf Gleichheit der Parametervektoren in den einzelnen Wellen.

Der Likelihood-Quotienten-Test (LR-Test; vgl. Abschnitt 2.3.2) mit der Nullhypothese

$$H_0 : \boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \ldots = \boldsymbol{\beta}_T = \boldsymbol{\beta},$$

mit $K = 11$ und mit $T = 13$, kann wie folgt formuliert werden:

$$\text{LR} = -2\, ln\, \lambda = 2 \cdot (ln\, \hat{L}_U - ln\, \hat{L}_R),$$

mit dem unrestringierten (restringierten) logarithmierten Likelihood-Schätzwert $ln\, \hat{L}_U$ ($ln\, \hat{L}_R$).

Im vorliegenden Fall erhält man als Schätzwert für die unrestringierte logarithmierte Likelihood, $ln\, \hat{L}_U$, (als Summe der dreizehn logarithmierten Likelihoods, jeweils ausgewertet an der Stelle des zugehörigen ML-Schätzers) $ln\, \hat{L}_U = -766.456$ (vgl. Tabellen 6.1.1 und 6.1.2) und als Schätzwert für die restringierte logarithmierte Likelihood, $ln\, \hat{L}_R$, (als Wert der logarithmierten Likelihoodfunktion des statischen Probitmodells über alle Wellen ausgewertet an der Stelle des zugehörigen ML-Schätzers) $ln\, \hat{L}_R = -803.999$ (vgl. Tabelle 6.2.2). Dieses statische Probitmodell über alle Wellen (vgl. Abschnitt 6.2.1) unterstellt implizit eine Gleichheit der einzelnen Parametervektoren $\boldsymbol{\beta}_t, t = 1,\ldots,13$.

Der Wert der Likelihood-Quotienten-Teststatistik ergibt sich somit zu:

$$LR = -2\, ln\, \boldsymbol{\lambda} = 2 \cdot (ln\, \hat{L}_U - ln\, \hat{L}_R) = 2 \cdot (-766.456 + 803.999) = 75.086.$$

Die Parametervektoren $\boldsymbol{\beta}_t, t = 1,\ldots,T$ enthalten jeweils elf Elemente; damit setzt sich die obige Nullhypothese aus insgesamt 132 $(12 \cdot 11)$ Restriktionen zusammen. Die LR-Teststatistik genügt demnach einer χ^2-Verteilung mit 132 Freiheitsgraden. Für das 0.95-Quantil der χ^2 - Verteilung mit 132 Freiheitsgraden ergibt sich ein Wert von 159.81.

Damit kann die oben angegebene Nullhypothese bei der unterstellten Irrtumswahrscheinlichkeit nicht abgelehnt werden; die Verwendung des der Schätzung der Parameter im Abschnitt 6.2.1 zugrundeliegenden Panelmodells anstelle der T Querschnittsmodelle führt zu keiner signifikanten Änderung der Werte der logarithmierten Likelihood. Somit steht das erhaltene Testergebnis in keinem Widerspruch zur Annahme gleicher Parametervektoren über alle Wellen.

6.3 Ein statisches Probitmodell mit zufälligen, individuenspezifischen Effekten unter der Annahme äquikorrelierter Störgrößen

Im statischen Probitmodell über die Wellen B bis N (vgl. Abschnitt 6.2.1) wurde davon ausgegangen, daß die Störgrößen (im Schwellenwertmodell) keine Korrelation über die Zeit aufweisen, d.h., man unterstellte voneinander unabhängige individuelle Entscheidungen über die Zeit (Annahme vollkommener Homogenität). Diese Annahme wird im statischen Probitmodell mit äquikorrelierten Störgrößen fallengelassen, indem stochastische, individuenspezifische Effekte explizit im Modell berücksichtigt werden (vgl. Abschnitt 2.2.2.1). Die quantitative Analyse dieses Modells zeigt zunächst folgendes (vgl. Tabelle 6.3.1):

– Der Schätzwert für die Varianz der zufälligen, individuenspezifischen Effekte $\alpha_i, i = 1,\ldots,N$ ist offenbar signifikant von null verschieden, d.h., die Nullhypothese $H_0\colon \sigma_\alpha^2 = 0$ kann (mit geringer Irrtumswahrscheinlichkeit) verworfen werden. Eine Parameterschätzung mittels des vollkommene Unabhängigkeit der individuellen Entscheidungen über die Wellen B bis N unterstellenden Probitmodells steht im vorliegenden Fall und aufgrund dieses statistischen Kriteriums offenbar im Widerspruch zum Datenbefund; die Parameterschätzung mittels jenes in Abschnitt 6.2 verwendeten Modells führt somit zu inkonsistenten Schätzwerten.

Bei expliziter Berücksichtigung stochastischer, individuenspezifischer Effekte im statischen Probitmodell ergeben sich vergleichsweise folgende Schätzergebnisse für die Koeffizienten der einzelnen Variablen (Tabelle 6.3.1; vgl. dazu Tabelle 6.2.2):

– Fast sämtliche Koeffizientenschätzwerte sind dem Betrag nach größer als die der Modellschätzung unter vollkommener Homogenitätsannahme. Gleichzeitig sind auch fast sämtliche asymptotische Standardfehler unter dieser Spezifikation größer als die ermittelten Standardfehler unter der Spezifikation in Abschnitt 6.2; dies ist eine Folge der unkorrekten Ermittlung dieser Schätzwerte im Vergleichsmodell.

– Der Einfluß der Altersvariablen ALT und ALTSQ auf die Wahrscheinlichkeit, arbeitslos zu werden, ist stärker ausgeprägt als im Vergleichsmodell. Durch Einsetzen der entsprechenden Werte der Variablen ALT und ALTSQ in die Funktion $f(\mathrm{ALT},\mathrm{ALTSQ}) = \hat{\boldsymbol{\beta}}_{\mathrm{ALT}} \cdot \mathrm{ALT} + \hat{\boldsymbol{\beta}}_{\mathrm{ALTSQ}} \cdot \mathrm{ALTSQ}$ ermittelt man hier als lokales Minimum ein Alter von 48 Jahren, d.h., das lokale Minimum bleibt annähernd unverändert.

– Der Einfluß der Variablen ′Schwerbehinderung′ (GZ) auf die Chance, arbeitslos zu werden, ist weniger gesichert nachweisbar als im Vergleichsmodell. Schwerbehinderte zeigen gegenüber nicht behinderten Personenen jedoch auch hier das signifikant höhere Risiko, arbeitslos zu werden.

Tabelle 6.3.1

Statisches Probitmodell mit zufälligen, individuenspezifischen Effekten unter der Annahme äquikorrelierter Störgrößen

Ergebnisse der Parameterschätzung

Wellen B (1985) bis N (1997)

Bezeichnung der Variablen	Parameter-Schätzwert	Asymptotischer t-Wert
ALT	-0.1658	-10.220
ALTSQ	0.1739	8.118
GZ	0.3318	2.142
BB1	-0.2096	-0.762
BST2	-1.2486	-9.501
NAT	0.2366	1.301
FST	-0.1280	-0.968
CHILD	-0.0240	-0.491
HINC	0.4661	13.810
ALQ	0.0789	3.809
ALD24	0.1199	15.590
σ_α^2	0.5233	5.986
LOG-L	-716.673	

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Es ist σ_α^2 die Varianz der zufälligen, individuenspezifischen Effekte und LOG-L der Wert der logarithmierten Likelihoodfunktion ausgewertet beim Maximum der Funktion.

Quelle: Eigene Berechnungen

- Ein Einfluß einer abgeschlossenen Berufsausbildung (BB1) ist auch im statischen Probitmodell mit äquikorrelierten Störgrößen signifikant nicht nachweisbar.

- Der Einfluß der Art des Arbeitsverhältnisses ′Angestellter/Arbeiter′ (BST2) auf die Beschäftigungswahrscheinlichkeit ist größer als im Vergleichsmodell und unter den getroffenen Annahmen noch gesicherter nachweisbar.

- Der Einfluß der Variablen ′Nationalität′ (NAT) auf die Wahrscheinlichkeit, arbeitslos zu werden, hat sich im Vergleich zum statischen Probitmodell tendenziell umgekehrt, ist aber auch hier signifikant nicht nachweisbar.

- Ein Einfluß der Variable ′Familienstand′ (FST) ist im Probitmodell mit äquikorrelierten Störgrößen, im Gegensatz zum vorher betrachteten Modell, nicht mehr signifikant nachweisbar.

- Auch die Anzahl der Kinder (CHILD) besitzt, folgt man dieser Schätzung, keinen signifikanten Einfluß auf die Wahrscheinlichkeit, beschäftigt zu sein.

- Der Einfluß des Nettohaushaltseinkommens (HINC) auf die Wahrscheinlichkeit, beschäftigt zu sein, ist im vorliegenden Probitmodell vergleichsweise größer und weiterhin sehr gesichert nachweisbar.

- Dem makroökonomischen Umfeld kommt weiterhin einige Bedeutung zu; so zeigt die Variable ′Arbeitslosenquote′ (ALQ) auch im statischen Probitmodell mit äquikorrelierten Störgrößen einen signifikant positiven Einfluß auf das Risiko, arbeitslos zu werden.

- Der Einfluß der Arbeitslosigkeitsdauer (ALD24) besitzt hier einen vergleichsweise geringeren positiven Einfluß auf die Wahrscheinlichkeit, arbeitslos zu werden, ist aber weiterhin sehr gesichert nachweisbar.

Zusammenfassend lassen sich für das statische Probitmodell (über die Wellen B bis N) mit äquikorrelierten Störgrößen die folgenden Aussagen treffen:

- Die Annahme vollkommener Unabhängigkeit der individuellen Entscheidungen über alle Wellen scheint empirisch nicht gerechtfertigt. Bei Unterstellung dieser Eigenschaften im statischen Probitmodell in Abschnitt 6.2 geht man offenbar von falschen Annahmen aus, was zu verzerrten und inkonsistenten Parameterschätzwerten führen kann.

- Der Wert der logarithmierten Likelihoodfunktion an der Stelle des ML-Schätzers wird durch Berücksichtigung stochastischer, individuenspezifischer Effekte im statischen Probitmodell erhöht (von -804 auf -717; vgl. Tabellen 6.2.2 und 6.3.1).

– Der Parameterschätzwert für die Variable ′Arbeitslosigkeitsdauer′ (ALD24) und dessen zugehöriger asymptotischer Schätzwert sind gegenüber dem vorher betrachteten Modell deutlich gesunken. Dieses Ergebnis unterstreicht die These, daß der Parameterschätzwert der Variablen ALD24 im Probitmodell des Abschnitts 6.2 verzerrt geschätzt wird, denn es liegt offenbar keine Unkorreliertheit der Störgrößen über die Zeit vor. Vielmehr existiert eine positive Korrelation zwischen der Variablen ALD24 und der Störgröße des Schwellenwertmodells, was im vorher betrachteten Modell zu einer Überschätzung des zugehörigen Parameterschätzwertes führte.

– Der Parameterschätzwert für die Varianz der stochastischen, individuenspezifischen Effekte, σ_α^2, der das Ausmaß der unbeobachteten Heterogenität im Modell wiederspiegelt, und formal definiert ist als die Korrelation der latenten endogenen Variablen von einer Welle zur nächsten, ist mit einem Wert von etwa 0.5 vergleichsweise hoch. Diese Tatsache deutet darauf hin, daß einige nicht beobachtete bzw. nicht meßbare Variable, oder bisher nicht im Modell berücksichtigte Variable, wie möglicherweise der um eine Periode verzögerte Arbeitsstatus, einen Einfluß auf die Beschäftigungswahrscheinlichkeit ausüben.

Dies spricht dafür, das statische Probitmodell mit äquikorrelierten Störgrößen durch die Berücksichtigung der Variablen ′Arbeitsstatus der Vorperiode′ (ESTL), als zusätzliche exogene Variable, zu modifizieren bzw. zu dynamisieren.

Bevor im Abschnitt 6.5 ein dynamisches Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand betrachtet wird, befaßt sich der folgende Abschnitt mit der quantitativen Analyse eines Probitmodells mit autokorrelierten Störgrößen unter Verwendung des MECOSA-Ansatzes.

6.4 Ein statisches Probitmodell unter der Annahme autokorrelierter Störgrößen

Im statischen Probitmodell mit äquikorrelierten Störgrößen (vgl. Abschnitt 6.3) wurde von einer über alle Wellen konstanten Korrelation der Störgrößen bzw. der latenten Zufallsvariablen ausgegangen. Im hier verwendeten statischen Probitmodell mit autokorrelierten Störgrößen wird dagegen unterstellt, daß die Störgrößen im Schwellenwertmodell einem Markov-Prozeß erster Ordnung folgen, d.h., es wird eine mit dem Faktor ρ abnehmende Korrelation der Störgrößen über die Zeit angenommen (vgl. Abschnitt 3.2.1). Diese Hypothese einer über die Zeit abnehmenden Abhängigkeit der individuellen Entscheidungen von Entscheidungen vorausgegangener Perioden, soll im Probitmodell mit autokorrelierten Störgrößen mit Hilfe des MECOSA-Ansatzes (vgl. Abschnitt 4.1) getestet werden.

Bei der Verwendung des MECOSA-Algorithmus aufgetretene Lösbarkeitsprobleme machen jedoch eine Modifikation des Panel-Zeitraums notwendig. Dieser wird auf die Wellen E bis N verkürzt. Im Gegensatz zu den vorher betrachteten Modellen findet die Variable ′Arbeitslosigkeitsdauer′ (ALD24) bei der empirischen Analyse des Probitmodells mit autokorrelierten Störgrößen keine Berücksichtigung. Dies hat folgenden Grund:

Im statischen Probitmodell mit äquikorrelierten Störgrößen hat sich gezeigt, daß die Varianz der stochastischen, individuenspezifischen Effekte signifikant von null verschieden ist, d.h., es ist nicht davon auszugehen, daß die Störgrößen im Schwellenwertmodell nicht korreliert sind. Diese Tatsache führt bei der empirischen Analyse der Probitmodelle einzelner Wellen zu einer verzerrten und inkonsistenten Schätzung des Parameters für die Variable ′Arbeitslosigkeitsdauer′ (ALD24), weil im vorliegenden Fall eine positive Korrelation zwischen dieser Variablen und der zugehörigen Störgröße des Schwellenwertmodells besteht. Da der MECOSA-Ansatz auf den Parameterschätzergebnissen einzelner Wellen aufbaut, erscheint es nicht sinnvoll, die Variable ALD24 bei der empirischen Analyse des Probitmodells mit autokorrelierten Störgrößen zu berücksichtigen (vgl. dazu auch die entsprechenden Ausführungen in Abschnitt 6.1, insbesondere die Fußnote auf S. 57).

Wie in den theoretischen Ausführungen zu diesem Ansatz behandelt (vgl. Abschnitt 4.1) ′startet′ der MECOSA-Algorithmus mit Querschnittsschätzungen der Parameter für die einzelnen Wellen (Stufe 1).

In der zweiten Stufe des MECOSA-Ansatzes erfolgt eine Schätzung der $\frac{T(T-1)}{2}$ Kovarianzen[1] für die Störgrößen der Wellen E bis N ($T = 10$; Tabelle 6.4.1).

Die Ergebnisse dieser Schätzung zeigen Kovarianzfolgen, die überwiegend nicht im Widerspruch zu der Hypothese mit einem Faktor ρ über die Zeit abnehmender Zusammenhänge stehen und damit nicht im Widerspruch zu der Hypothese abnehmender Abhängigkeit der individuellen Entscheidungen über die Zeit.
Andererseits werden hier vergleichsweise hohe Kovarianzen für zeitlich nicht weit von einander entfernte Wellen ermittelt; dies spricht für die These, daß im hier betrachteten Modell wichtige Einflußvariablen nicht berücksichtigt wurden bzw. nicht berücksichtigt werden konnten und damit für eine Dynamisierung des Modells.

In der dritten Stufe der MECOSA-Schätzung wird getestet, ob der diese zeitliche Entwicklung der Kovarianzen charakterisierende Korrelationskoeffizient ρ signifikant von null verschie-

---

[1] Im vorliegenden Fall sind, aufgrund der Normierung, die Kovarianzen und die entsprechenden Korrelationen identisch.

Tabelle 6.4.1: **Statisches Probitmodell unter der Annahme autokorrelierter Störgrößen**

Ergebnisse der MECOSA-Schätzung der Kovarianzmatrix der Störgrößen

Wellen E (1988) bis N (1997)

Stufe 2

	WELLE E	WELLE F	WELLE G	WELLE H	WELLE I	WELLE J	WELLE K	WELLE L	WELLE M	WELLE N
WELLE E	1	0.877	0.808	0.905	0.747	0.709	0.757	0.356	0.494	0.319
WELLE F	0.877	1	0.902	0.839	0.695	0.387	0.714	0.327	0.376	0.303
WELLE G	0.808	0.902	1	0.912	0.908	0.817	0.804	0.695	0.466	0.434
WELLE H	0.905	0.839	0.912	1	0.893	0.785	0.730	0.539	0.534	0.372
WELLE I	0.747	0.695	0.908	0.893	1	0.952	0.852	0.876	0.652	0.648
WELLE J	0.709	0.387	0.817	0.785	0.952	1	0.938	0.784	0.830	0.729
WELLE K	0.757	0.714	0.804	0.730	0.852	0.938	1	0.921	0.769	0.839
WELLE L	0.356	0.327	0.695	0.539	0.876	0.784	0.921	1	0.814	0.764
WELLE M	0.494	0.376	0.466	0.534	0.652	0.830	0.769	0.814	1	0.955
WELLE N	0.319	0.303	0.434	0.372	0.648	0.729	0.839	0.764	0.955	1

Quelle: Eigene Berechnungen; aufgrund der Normierung der Varianzen der Störgrößen ist dies zugleich die Korrelationsmatrix.

Tabelle 6.4.2

Statisches Probitmodell unter der Annahme autokorrelierter Störgrößen

Ergebnisse der Parameterschätzung für das Panel unter Verwendung des MECOSA-Algorithmus

Wellen E (1988) bis N (1997)

Stufe 3

Variable	Parameter-Schätzwert	Asympt. t-Wert
ALT	-0.1213	-0.8661
ALTSQ	0.1934	1.1926
GZ	0.8296	3.1444
BB1	-0.2618	-1.0149
BST2	-0.7794	-1.9304
NAT	-0.3806	-1.8697
FST	-0.8198	-3.6885
CHILD	-0.0071	-0.0773
HINC	0.2017	5.7381
ALQ	0.2976	0.4099
σ_α^2	0.4139	0.7216
ρ	0.9184	10.8806

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Zur Definition der Parameter σ_α^2 und ρ vgl. Abschnitt 3.2.1.

Quelle: Eigene Berechnungen

den ist oder nicht. Zugleich erfolgt hier eine Schätzung der strukturellen Parameter des Panelmodells unter Berücksichtigung der angenommenen Korrelationsstruktur (Tabelle 6.4.2):

- Der Schätzwert für den Korrelationskoeffizienten ρ ist, wie zu erwarten, signifikant von null verschieden. Dieses Ergebnis stützt die Annahme einer (überwiegend) abnehmenden Abhängigkeit der individuellen Entscheidungen über die Zeit.[1]

- Der Parameterschätzwert für σ_α^2 ist nicht signifikant von null verschieden.

- Faßt man die Ergebnisse der Schätzung der Koeffizienten zusammen, so ist festzustellen, daß dieses Modell hier weniger zur Beschreibung individueller Arbeitslosigkeit geeignet ist, als das in Abschnitt 6.3 beschriebene Modell. Für die Schätzwerte der strukturellen Koeffizienten ermittelt man durchweg größere (asymptotische) Standardfehler, was höchstens zu einem vernachlässigbaren Teil auf die geringere Anzahl von Freiheitsgraden zurückgeführt werden kann.

Das wesentliche Ergebnis der Verwendung des MECOSA-Ansatzes bei einer Analyse individueller Arbeitslosigkeit ist also darin zu sehen, daß die Schätzergebnisse eine mit dem Faktor ρ signifikant abnehmende Abhängigkeit der individuellen Entscheidungen über die Zeit nahelegen.

Dieses Ergebnis kann seine Ursache auch in der Nichtberücksichtigung der Variablen ′Arbeitsstatus der Vorperiode′ (ESTL) und ′Arbeitslosigkeitsdauer′ (ALD24) haben und somit ein Hinweis auf eine notwendige Dynamisierung des Modells sein; es unterstreicht damit die Notwendigkeit, bei einer Analyse individueller Arbeitslosigkeit sowohl autokorrelierte Störgrößen, als auch verzögert endogene Variablen im jeweiligen Probitmodell (Schwellenwertmodell) zuzulassen. Dynamische Probitmodelle mit äquikorrelierten bzw. autokorrelierten Störgrößen werden in den folgenden Abschnitten betrachtet.

---

[1] Eine Schätzung des statischen Probitmodells des Abschnitts 6.3 über die Wellen E bis N und ohne die Berücksichtigung der Variablen ALD24 (entsprechend der empirischen Analyse des statischen Probitmodells mit autokorrelierten Störgrößen) liefert für den Korrelationskoeffizienten σ_α^2 einen Wert von 0.5727 bei einem asymptotischen t-Wert von 4.3173; daraus errechnet sich ein asymptotisches Konfidenzintervall zum Konfidenzniveau $1-\boldsymbol{\alpha}=0.99$ für σ_α^2 von KI = [0.2311 ; 0.9143]. Man erkennt somit, daß eine signifikante Äquikorrelation der Störgrößen nicht zwingend im Widerspruch zu der Annahme autokorrelierter Störgrößen steht; alle ermittelten Kovarianzen der Stufe 2 des MECOSA-Ansatzes (vgl. Tabelle 6.4.1) liegen innerhalb des obigen Konfidenzintervalls.

6.5 Ein dynamisches Probitmodell unter der Annahme äquikorrelierter Störgrößen - Exogener Anfangszustand

Die Einbeziehung der Variablen ′Arbeitsstatus der Vorperiode′ (ESTL) als erklärende Variable in ein statisches Probitmodell mit äquikorrelierten Störgrößen, führt zu einem dynamischen Probitmodell mit äquikorrelierten Störgrößen (vgl. Abschnitt 3.2.2). Zusätzlich wird hier vereinfachend die Annahme eines exogenen Anfangszustandes getroffen. Die empirische Analyse unter Annahme dieses Probitmodells liefert die in Tabelle 6.5.1 ausgewiesenen Schätzwerte. Diese Schätzwerte lassen - im Vergleich zu den Schätzwerten für das entsprechende statische Probitmodell (vgl. Tabelle 6.3.1) - folgende Schlußfolgerungen zu:

- Die asymptotischen Standardfehler der Koeffizientenschätzwerte sind weitgehend unverändert.

- Der Koeffizientenschätzwert der Variablen ′Arbeitsstatus der Vorperiode′ (ESTL) legt eine signifikante Zustandsabhängigkeit nahe, d.h., die Arbeitslosigkeit einer Person in der Vorperiode erhöht offenbar signifikant deren Risiko, ′heute′ arbeitslos zu sein. Gleichzeitig sinkt der Einfluß der Varianz der stochastischen, individuenspezifischen Effekte, σ_α^2, d.h., ein Teil der ′unbeobachteten Heterogenität′ wird durch die Variable ESTL erfaßt. Dieses Ergebnis unterstreicht die Notwendigkeit einer Dynamisierung, zeigt aber zugleich, daß trotz dieser Dynamisierung des Modells eine signifikante Abhängigkeit der individuellen Entscheidungen über die Zeit verbleibt.

- Der Einfluß der Altersvariablen ALT und ALTSQ ist gegenüber dem entsprechenden statischen Probitmodell nahezu unverändert. Das Arbeitslosigkeitsrisiko sinkt zunächst mit zunehmendem Alter; das Risiko nimmt aber ab einem Alter der betrachteten Personen von 46 Jahren wieder zu.

- Im Gegensatz zum entsprechenden statischen Probitmodell mit äquikorrelierten Störgrößen ist der Effekt einer Schwerbehinderung (GZ) auf die Wahrscheinlichkeit, beschäftigt zu sein, abgeschwächt und nicht mehr signifikant nachweisbar.

- Ein Einfluß einer abgeschlossenen Berufsausbildung (BB1) ist auch hier signifikant nicht nachweisbar.

- Die Art des Beschäftigungsverhältnisses, ′Angestellter/Arbeiter′ (BST2), besitzt weiterhin einen signifikanten, die Chance arbeitslos zu werden beeinflussenden Effekt, d.h., Angestellte zeigen (im Vergleich zu Arbeitern) auch in diesem Modell ein signifikant niedrigeres Risiko, arbeitslos zu werden.

Tabelle 6.5.1

Dynamisches Probitmodell unter der Annahme äquikorrelierter Störgrößen - Exogener Anfangszustand

Ergebnisse der Parameterschätzung

Wellen B (1985) bis N (1997)

Bezeichnung der Variablen	Parameter-Schätzwert	Asymptotischer t-Wert
ALT	-0.1528	-10.610
ALTSQ	0.1649	8.501
GZ	0.2582	1.585
BB1	-0.0718	-0.304
BST2	-1.1623	-8.315
NAT	0.0294	0.183
FST	-0.2177	-1.741
CHILD	-0.0317	-0.637
HINC	0.4072	11.842
ALQ	0.0672	3.237
ALD24	0.0932	11.401
ESTL	0.8757	7.634
σ_α^2	0.3817	4.267
LOG-L	-697.283	

Legende: Zur Variablenbezeichnung vgl. S. 48f. - Es ist σ_α^2 die Varianz der zufälligen, individuenspezifischen Effekte und LOG-L der Wert der logarithmierten Likelihoodfunktion ausgewertet beim Maximum der Funktion.

Quelle: Eigene Berechnungen

- Ein Einfluß der Nationalität (NAT) auf die Beschäftigungswahrscheinlichkeit ist bei dieser Spezifikation nicht nachweisbar.

- Bei gleicher Wirkungsrichtung ist das Ausmaß des Einflusses des Familienstandes (FST) auf das Arbeitslosigkeitsrisiko erhöht.

- Der Einfluß der Kinderzahl (CHILD) auf das Risiko, arbeitslos zu werden, ist auch hier nicht nachweisbar.

- Der positive Einfluß des Nettohaushaltseinkommens (HINC) auf das Risiko, arbeitslos zu werden, ist - im Vergleich zu dem entsprechenden statischen Probitmodell - etwas geringer.

- Die Dauer der Arbeitslosigkeit in den zurückliegenden zwei Jahren (ALD24) zeigt auch hier einen - wenn auch leicht abgeschwächten - signifikant positiven Einfluß auf das Arbeitslosigkeitsrisiko der betrachteten Person.

Zusammenfassend lassen sich für das dynamische Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand folgende wichtige Punkte festhalten:

- Die empirische Analyse des dynamischen Probitmodells mit äquikorrelierten Störgrößen und exogenem Anfangszustand zeigt die Bedeutung einer Dynamisierung der statischen Probitmodelle bei der Analyse individueller Arbeitslosigkeit. Für die Dynamisierung sprechen insbesondere der gesichert nachweisbare Einfluß der verzögert endogenen Variable ESTL auf die Beschäftigungswahrscheinlichkeit, d.h., eine vorliegende Zustandsabhängigkeit, und die Abnahme des Schätzwertes für σ_α^2.

- Der weiterhin signifikant von null verschiedene Parameterschätzwert für σ_α^2, die Varianz der stochastischen, individuenspezifischen Effekte, deutet darauf hin, daß auch dieses Modell möglicherweise unvollständig ist, daß nicht im Modell berücksichtigte Variable oder nicht meßbare Variable, wie beispielsweise die Motivation (zu arbeiten) oder bestimmte Fähigkeiten der betrachteten Person, einen Einfluß auf die Beschäftigungswahrscheinlichkeit ausüben. Diese Tatsache unterstreicht die Bedeutung der hier vorgenommenen Berücksichtigung stochastischer, individuenspezifischer Effekte im zugehörigen Schwellenwertmodell (Annahme unbeobachteter Heterogenität).

Im folgenden Abschnitt wird die Schätzung der marginalen Effekte (in nichtlinearen Panelmodellen) betrachtet; dieses geschieht beispielhaft für das dynamische Probitmodell mit äquikorrelierten Störgrößen und als exogen angenommenem Anfangszustand.

6.6 Ermittlung marginaler Effekte für ausgewählte Probitmodelle

Bei der empirischen Analyse statischer und dynamischer Probitmodelle wurde in der vorliegenden Arbeit bisher die Schätzung marginaler Effekte vernachlässigt. Dies soll in diesem Abschnitt beispielhaft für einige der Probitmodelle (vgl. die Tabellen 6.2.2, 6.3.1 und 6.5.1) nachgeholt und für einen Fall diskutiert werden. Dabei wird aufgezeigt, daß die Schätzung der marginalen Effekte für die vorliegende Panelstichprobe zu wenig aussagekräftigen Ergebnissen führt.

Die marginalen Effekte liefern die Information, um welchen Wert sich die Wahrscheinlichkeit, beschäftigt zu sein, ändert, wenn sich eine bestimmte exogene Variable um eine Einheit erhöht. Für die in der vorliegenden Arbeit betrachteten Probitmodelle ergeben sich - wie oben (vgl. S. 24) angeführt - die marginalen Effekte zu:

$$\frac{\partial\, P(y_{it}=1)}{\partial\, \boldsymbol{x_{it}}} = \frac{\partial\, \boldsymbol{\Phi}(\boldsymbol{x}_{it}'\boldsymbol{\beta})}{\partial\, \boldsymbol{x_{it}}} = \boldsymbol{\varphi}(\boldsymbol{x}_{it}'\boldsymbol{\beta}) \cdot \boldsymbol{\beta}\,,$$

mit $\boldsymbol{\Phi}$, der Verteilungsfunktion und $\boldsymbol{\varphi}$, der Dichtefunktion der Standardnormalverteilung.

Aus dieser Gleichung kann man entnehmen, daß die marginalen Effekte von dem Strukturparametervektor $\boldsymbol{\beta}$ und dem Vektor der exogenen Variablen, $\boldsymbol{x_{it}}$, abhängen. Zur Schätzung der marginalen Effekte wird deshalb in der Regel der Vektor der Mittelwerte, $\bar{\boldsymbol{x}} = \frac{1}{N \cdot T}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{x_{it}}$, herangezogen, womit sich die geschätzten marginalen Effekte zu $\boldsymbol{\varphi}(\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}}) \cdot \hat{\boldsymbol{\beta}}$ ergeben.

Die Schätzwerte der marginalen Effekte für ausgewählte Probitmodelle und deren zugehörige asymptotische *t*-Werte sind in der Tabelle 6.6.1 zusammengefaßt. Die Diskussion dieser Schätzergebnisse wird im folgenden beispielhaft für das dynamische Probitmodell unter der Annahme äquikorrelierter Störgrößen und eines als exogen angenommenen Anfangszustandes geführt (vgl. die Spalten 5 und 6 der Tabelle 6.6.1).

Die Dichtefunktion der Standardnormalverteilung an der Stelle $\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}}$, $\boldsymbol{\varphi}(\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}})$, nimmt für das dynamische Probitmodell den Wert von 0.0069912 an, da $\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}} = -2.844$. Die Schätzwerte der marginalen Effekte in diesem Modell ergeben sich also dadurch, daß man die Parameterschätzwerte des dynamischen Probitmodells mit dem Wert der Dichtefunktion an der Stelle $\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}} = -2.844$ multipliziert (zur Ermittlung der asymptotischen Standardfehler vgl. Greene, 1997, S. 884f.).

Die Verteilungsfunktion der Standardnormalverteilung liefert an der Stelle $\bar{\boldsymbol{x}}'\hat{\boldsymbol{\beta}}$ einen Wert von 0.0022276. Betrachtet man also den Mittelwertvektor $\bar{\boldsymbol{x}}$ als die Realisation des Vektors

Tabelle 6.6.1

Ergebnisse der Schätzung der marginalen Effekte

Ausgewählte Panelmodelle

Wellen B (1985) bis N (1997)

	Statisches Probitmodell (vgl. Tabelle 6.2.2)		Statisches Probitmodell Zufällige Effekte Äquikorrelation (vgl. Tabelle 6.3.1)		Dynamisches Probitmodell Zufällige Effekte Äquikorrelation Exogener Anfangszustand (vgl. Tabelle 6.5.1)	
Bezeichnung der Variablen	Marginaler Effekt	Asympt. t-Wert	Marginaler Effekt	Asympt. t-Wert	Marginaler Effekt	Asympt. t-Wert
	1	2	3	4	5	6
ALT	-0.004031	-7.730	-0.000546	-2.394	-0.001070	-2.802
ALTSQ	0.004480	6.872	0.000573	2.320	0.001155	2.688
GZ	0.013527	3.504	0.001093	1.733	0.001810	1.460
BB1	-0.004339	-1.002	-0.000691	-0.760	-0.000503	-0.307
BST2	-0.035341	-9.334	-0.004113	-2.442	-0.008139	-2.930
NAT	-0.005369	-1.454	0.000779	1.287	0.000206	0.185
FST	-0.008554	-2.467	-0.000422	-0.935	-0.001525	-1.546
CHILD	0.000185	0.177	-0.000079	-0.485	-0.000222	-0.638
HINC	0.009045	8.498	0.001535	2.530	0.002852	3.088
ALQ	0.001672	3.247	0.000260	2.111	0.000471	2.246
ALD24	0.005021	8.599	0.000395	2.286	0.000652	2.581
ESTL					0.006133	2.604

Legende: Zur Variablenbezeichnung vgl. S. 48f.

Quelle: Eigene Berechnungen

der exogenen Variablen einer ′typischen′ Person, so nimmt deren Wahrscheinlichkeit, arbeitslos zu werden, den Wert 0.0022276 an.
In diesem Zusammenhang ist es wichtig festzuhalten, daß das prognostizierte Arbeitslosigkeitsrisiko für den Mittelwertvektor der exogenen Variablen $\bar{x}$ ($\Phi(\bar{x}'\hat{\beta})$), im vorliegenden Fall einer sehr ungleich besetzten Panelstichprobe, in der Regel stark vom durchschnittlich prognostizierten Arbeitslosigkeitsrisiko (dem tatsächlich in der Panelstichprobe existierenden Arbeitslosigkeitsrisiko) abweicht. Dieses Risiko ist definiert als $\frac{1}{N \cdot T}\sum_{i=1}^{N}\sum_{t=1}^{T}\Phi(x_{it}'\hat{\beta})$ und liefert für die vorliegende Panelstichprobe einen Wert von 0.0547; dieser Wert überschreitet das prognostizierte Arbeitslosigkeitsrisiko für den Mittelwertvektor der exogenen Variablen im vorliegenden Fall um mehr als das Zwanzigfache. Diese Tatsache unterstreicht die geringe Aussagekraft marginaler Effekte in sehr ungleich besetzten Antwortmodellen (Probitmodellen).

Der Schätzer für den marginalen Effekt der Variablen ′Nettohaushaltseinkommen′ (HINC) liefert im dynamischen Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand beispielsweise einen Wert von 0.00285, d.h., steigt das Nettoeinkommen des Haushaltes einer betrachteten Person um 1000 DM, so erhöht sich dessen Arbeitslosigkeitsrisiko um 0.285 Prozent. Für eine Person, die in der Vorperiode arbeitslos war steigt dieses Risiko im dynamischen Probitmodell mit äquikorrelierten Störgrößen beispielsweise um 0.613 Prozent. Die Tatsache, daß sich das prognostizierte Arbeitslosigkeitsrisiko (mit einem Wert von 0.23 Prozent) um 0.613 Prozent erhöht, falls die betrachtete Person in der Vorperiode arbeitslos war, besitzt nur eine sehr geringe Aussagekraft.

Neben dem Problem einer sehr ungleich besetzten Panelstichprobe gilt es bei der Schätzung der marginalen Effekte in den hier betrachteten Probitmodellen noch folgenden Punkt zu beachten: Die Berechnung bzw. die Schätzung der marginalen Effekte durch die Differentiation der Normalverteilungsfunktion (an der Stelle $x_{it}'\beta$) nach dem Vektor der exogenen Variable $\mathbf{x}_{it}$ ist in der Regel nur für stetige Variablen sinnvoll (vgl. Greene, 1997, S.878f.). Da die hier betrachtete Panelstichprobe viele binäre exogene Variable enthält, erscheint die obige Vorgehensweise bei der Schätzung der marginalen Effekte auch unter diesem Gesichtspunkt als nicht sinnvoll.

Zum Abschluß des sechsten Kapitels wird im folgenden Abschnitt die Güte der Anpassung in ausgewählten Probitmodellen betrachtet.

6.7 Güte der Anpassung in ausgewählten Probitmodellen

Als Maß für die Güte der Anpassung bzw. als Maß für die Prognosefähigkeit der verschiedenen Probitmodelle werden Vierfeldertafeln betrachtet, deren Ausprägungen Auskunft darüber

geben, wie häufig die Beschäftigungsstatus (beschäftigt/arbeitslos) im jeweils betrachteten Panelmodell richtig bzw. falsch prognostiziert wurden.

Der prognostizierte Beschäftigungsstatus für das Individuum i zum Zeitpunkt t, $\hat{y}_{it}$, ergibt sich dabei in allen betrachteten Panelmodellen gemäß folgender (Schätz-) Regel:

$$\hat{y}_{it} = \begin{cases} 1 & \text{falls } \boldsymbol{\Phi}(\boldsymbol{x}_{it}'\hat{\boldsymbol{\beta}}) > 0.5 \\ 0 & \text{sonst.} \end{cases}$$

Die Diskussion der in der Tabelle 6.7.1 dargelegten (Schätz-) Ergebnisse wird im folgenden beispielhaft für das dynamische Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand geführt, da sich alle betrachteten Panelmodelle hinsichtlich ihrer Prognosefähigkeiten nur geringfügig unterscheiden:

- In den dreizehn Wellen des Panels sind die 527 Individuen in 6476 der Fälle beschäftigt und in 375 der Fälle arbeitslos.

- Der Beschäftigungsstatus 0 (′beschäftigt′) wird in 99.53 Prozent der Fälle richtig prognostiziert; dagegen wird der Beschäftigungsstatus 1 (′arbeitslos′) nur in 37.07 Prozent der Fälle richtig vorhergesagt. Insgesamt werden die Beschäftigungsstatus im betrachteten Panelmodell in 96.12 Prozent der Fälle korrekt vorhergesagt.

Diese scheinbar hohe Prognosegüte der Anpassung relativiert sich, wenn man bedenkt, daß eine ′naive′ No-change-Vorhersage, die in allen Fällen eine 0 für den Beschäftigungsstatus der Individuen vorhersagt, in 94.53 Prozent der Fälle eine richtige Prognose liefern würde. Diese Tatsache spricht jedoch nicht gegen die im dynamischen Probitmodell erhaltenen Parameterschätzwerte, sondern eher gegen das an dieser Stelle verwendete Maß für die Güte der Anpassung.

Insbesondere für ungleich besetzte Panelstichproben, in denen die endogene Variable wesentlich häufiger den Zustand 0 als den Zustand 1 annimmt (oder umgekehrt), erweist sich der (Schwellen-) Wert von 0.5 in der obigen Schätzregel als nicht gut geeignet. Da die durchschnittlich prognostizierte Wahrscheinlichkeit für den Beschäftigungsstatus 1 im dynamischen Probitmodell bei etwa fünf Prozent liegt, bedarf es einer extremen Ausprägung des Vektors der Regressoren, $\boldsymbol{x}_{it}$, um $\boldsymbol{\Phi}(\boldsymbol{x}_{it}'\hat{\boldsymbol{\beta}}) > 0.5$ zu erhalten. Diese Tatsache ist u.a. der Grund dafür, daß nur etwa ein Drittel der Arbeitslosigkeitszustände der Individuen (im betrachteten Modell) richtig prognostiziert wurde.

Abschließend soll in diesem Abschnitt noch darauf hingewiesen werden, daß die Parameterschätzung in allen hier betrachteten Panelmodellen nicht das Ziel verfolgte ein bestimmtes Maß für die Anpassungsgüte im jeweiligen Modell zu optimieren. Die Parameterschätzwerte

Tabelle 6.7.1

Ergebnisse der Schätzung für die Güte der Anpassung

Ausgewählte Panelmodelle

Wellen B (1985) bis N (1997)

	Statisches Probitmodell (vgl. Tabelle 6.2.2)			Statisches Probitmodell Zufällige Effekte Äquikorrelation (vgl. Tabelle 6.3.1)			Dynamisches Probitmodell Zufällige Effekte Äquikorrelation Exogener Anfangszustand (vgl. Tabelle 6.5.1)		
	Prognostizierter Arbeitsstatus			Prognostizierter Arbeitsstatus			Prognostizierter Arbeitsstatus		
Tatsächlicher Arbeitsstatus	0	1	Summe	0	1	Summe	0	1	Summe
0	6430	46	6476	6446	30	6476	6446	30	6476
1	221	154	375	252	123	375	236	139	375
Summe	6651	200	6851	6698	153	6851	6682	169	6851

Legende: Bei der Variablen Arbeitsstatus steht die 0 für ′beschäftigt′, die 1 für ′arbeitslos′.

Quelle: Eigene Berechnungen

wurden hingegen so ermittelt, daß sie im jeweiligen Probitmodell die zugehörige logarithmierte Likelihoodfunktion maximieren.

7 Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit Panelmodellen mit dichotomer endogener Variable. Die Formulierung der nichtlinearen Panelmodelle erfolgt jeweils auf der Grundlage eines Schwellenwertmodells. Dabei wird in allen betrachteten Modellen von (standard-) normalverteilten Störgrößen ausgegangen, d.h., es werden in allen für die empirische Analyse verwendeten Modellen Probitmodelle unterstellt.

Die Implementierung der in der vorliegenden Arbeit behandelten methodischen Ansätze erfolgt mittels ausgewählter Daten des Deutschen Sozio-ökonomischen Panels. Hierbei geht es um eine Analyse individueller Arbeitslosigkeit auf dem deutschen Arbeitsmarkt und zwar von männlichen Arbeitern und Angestellten. Das der empirischen Analyse zugrundeliegende Datenpanel umfaßt insgesamt 13 Wellen (1985 bis 1997) und 527 Personen.

Im Nachhinein erscheint es bei der empirischen Analyse individueller Arbeitslosigkeit sinnvoll, dynamische nichtlineare Panelmodelle (Probitmodelle) zu unterstellen, da von einem Effekt einer in der Vergangenheit bestehenden Arbeitslosigkeit auf den zukünftigen Arbeitsstatus der betrachteten Person ausgegangen werden muß. Andererseits gilt es durch die explizite Berücksichtigung stochastischer, individuenspezifischer Effekte im Schwellenwertmodell, eine in der Regel vorhandene Heterogenität (der betrachteten Personen) im jeweiligen Probitmodell abzubilden.

Die zu Beginn der quantitativen Analyse individueller Arbeitslosigkeit noch bestehende Unsicherheit bzgl. der Modellformulierung begründet die zugrundeliegende Vorgehensweise der hier vorgelegten Arbeit. Die Ansätze für die Parameterschätzung in den betrachteten Probitmodellen beruhen in allen Fällen auf der ML-Methode, wobei die Parameter der in den verschiedenen Modellen berücksichtigten Variablen in vielen Fällen mit Hilfe einer marginalen logarithmierten Likelihoodfunktion geschätzt werden.

Die empirische Analyse der individuellen Arbeitslosigkeit männlicher Arbeiter und Angestellter lieferte u.a. folgende Ergebnisse:

Zunächst wurde gezeigt, daß die Annahme eines über alle Wellen gleichen Strukturparametervektors $\boldsymbol{\beta}$, die implizit in allen betrachteten Panelmodellen getroffen wurde, bei Unterstellung üblicher Irrtumswahrscheinlichkeiten und eines statischen Probitmodells, nicht verworfen werden kann (Abschnitt 6.2.2).

Die anschließende Schätzung eines statischen Probitmodells mit äquikorrelierten Störgrößen zeigt dann, daß die Varianz der stochastischen, individuenspezifischen Effekte signifikant von null verschieden ist. Damit wurde der Nachweis erbracht, daß im vorliegenden Fall von unbeobachteter Heterogenität auszugehen ist. Eine Parameterschätzung mittels eines

statischen Probitmodells, welches von vollkommener Unabhängigkeit der individuellen Entscheidungen über die Zeit ausgeht, stünde damit im Widerspruch zum Datenbefund und wurde deshalb für den verwendeten Datensatz als für die empirische Analyse individueller Arbeitslosigkeit wenig geeignet eingestuft. Dieser Befund wurde untermauert durch den deutlichen Anstieg des Wertes der logarithmierten Likelihoodfunktion beim Übergang vom herkömmlichen statischen Probitmodell auf das statische Probitmodell mit äquikorrelierten Störgrößen.

Der Parameterschätzwert für die Varianz der stochastischen, individuenspezifischen Effekte im statischen Probitmodell deutete darauf hin, daß nicht im Modell berücksichtigte Variable, wie beispielsweise der um eine Periode verzögerte Arbeitsstatus, einen Einfluß auf die individuelle Arbeitslosigkeit ausüben, was zu einer entsprechenden Dynamisierung des betrachteten Modells führte.

Die zusätzliche Analyse eines Probitmodells mit autokorrelierten Störgrößen mittels des MECOSA-Ansatzes zeigte eine mit dem Faktor ρ signifikant abnehmende Abhängigkeit der individuellen Entscheidungen über die Zeit (dabei ist $\rho > 0.9$). Dies legte eine Betrachtung komplexerer Autokorrelationsprozesse der Störgrößen als der des Varianzkomponentenmodells und/oder eine Dynamisierung des Modells durch Einbeziehen des Arbeitsstatus der Vorperiode nahe.

Im dynamischen Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand zeigt sich eine signifikante Zustandsabhängigkeit, d.h., die Arbeitslosigkeit einer Person in der Vorperiode erhöht deren Risiko, in der aktuellen Periode arbeitslos zu sein, signifikant. Außerdem unterstreicht der gesunkene Parameterschätzwert für die Varianz der zufälligen, individuenspezifischen Effekte in diesem Modell die Notwendigkeit der Dynamisierung des vorher statischen Probitmodells.
Dieses Modell erweist sich unter allen in der vorliegenden Arbeit betrachteten Modellen, als das zur Abbildung individueller Arbeitslosigkeit am besten geeignete.

Die quantitative Analyse des dynamischen Probitmodells mit äquikorrelierten Störgrößen liefert dabei die wichtige Erkenntnis, daß eine erfolgreiche Bekämpfung der ′heutigen′ Arbeitslosigkeit, beispielsweise durch (Um-) Schulungsmaßnahmen, einen signifikant positiven Effekt auf die Höhe der zukünftigen Arbeitslosigkeit besitzt, da sowohl der Arbeitslosigkeitszustand in der Vergangenheit (′Arbeitsstatus der Vorperiode′ (ESTL)), als auch die Arbeitslosigkeitsdauer in der Vergangenheit (′Arbeitslosigkeitsdauer′ (ALD24)) einen sehr gesichert nachweisbaren negativen Effekt auf die individuelle Beschäftigungswahrscheinlichkeit besitzen. Diese Tatsache dürfte auch von politischem Interesse sein.

Ebenfalls von politischem Interesse dürfte die Tatsache sein, daß ein hohes ′Nettohaushaltseinkommen′ (HINC) möglicherweise das individuelle Arbeitslosigkeitsrisiko signifikant

erhöht, da es offensichtlich zu einem geringeren Arbeitsanreiz für die betrachteten Personen führt.

Bei der empirischen Analyse individueller Arbeitslosigkeit wäre es ′wünschenswert′ gewesen, ein dynamisches Probitmodell mit autokorrelierten Störgrößen zu betrachten; da dieses Modell mit den dem Verfasser bekannten Verfahren jedoch nicht konsistent schätzbar war, konnte es nicht verwendet werden.

Die Vorgehensweise bei der Abbildung individueller Arbeitslosigkeit in der vorliegenden Arbeit unterliegt insgesamt der Restriktion, daß die zu betrachtenden Modelle auch konsistent schätzbar sein müssen. Das dynamische Probitmodell mit äquikorrelierten Störgrößen und exogenem Anfangszustand stellt somit vielleicht nicht das am meisten ′gewünschte′ Modell zur quantitativen Analyse individueller Arbeitslosigkeit dar, aber es ist noch konsistent schätzbar.

Zum Abschluß der vorliegenden Arbeit muß betont werden, daß die sich in den verschiedenen Probitmodellen ergebenden Parameterschätzwerte zwar die jeweilige logarithmierte Likelihoodfunktion maximieren, daß diese ML-Schätzwerte jedoch nicht automatisch eine gute Prognosefähigkeit bzw. eine hohe Güte der Anpassung im jeweiligen Probitmodell (hinsichtlich der Arbeitsstatus der betrachteten Personen) garantieren. Für die zukünftige Forschung bleibt deshalb u.a. die Frage, ob es sinnvoller ist, die Parameterschätzwerte so zu bestimmen, daß sie die gemeinsame Dichtefunktion maximieren (ML-Schätzung) oder so, daß die endogenen Variablen des jeweiligen Modells möglichst gut prognostiziert werden.

8 Literatur

Abramowitz, M and Stegun, I. A. (1964), Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. Washington, DC: US Government Print Office.

Amemiya, T. (1981), Qualitative Response Models: A Survey. *Journal of Economic Literature* 19, 483-536.

Amemiya, T. (1985), Advanced Econometrics. Oxford: Basil Blackwell.

Andersen, E. B. (1973), Conditional Inference and Models for Measuring. Copenhagen: Mentalhygiejnisk Forlag.

Arminger, G. (1992), Analyzing Panel Data with Non-Metric Dependent Variables - Probit Models, Generalized Estimating Equations, Missing Data and Absorbing States. Deutsches Institut für Wirtschaftsforschung, Discussion Paper No. 59. Berlin.

Arminger, G. (1994), Probit-Models for the Analysis of Non-Metric Panel Data. *Allgemeines Statistisches Archiv,* 78, 125-140.

Balestra, P. (1992), Introduction to Linear Models for Panel Data. In: Matyas, L. and P. Sevestre, (eds.), The Econometrics of Panel Data. Dordrecht: Kluwer, 21-29.

Berndt, E., Hall, B., Hall, R. and Hausman, J. (1974), Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement,* 3/4, 653-665.

Bronstein, I. N. und Semendjajev, K. A. (1987), Taschenbuch der Mathematik. Thun und Frankfurt/Main: Harri Deutsch.

Butler, J. and R. Moffitt (1982), A Computationally Efficient Quadrature Procedure of the One-Factor Multinomial Probit Model. *Econometrica,* 50, 761-764.

Chamberlain, G. (1984), Panel Data. In Griliches, Z. and M. D. Intriligator, (eds.), Handbook of Econometrics. Vol. 2. Amsterdam: North Holland, 1247-1318.

Chamberlain, G. (1985), Hetrogeneity, Omitted Variable Bias, and Duration Dependence. In: Heckman, J. and B. Singer, (eds.), Longitudinal Analysis of Labour Market Data. Cambridge University Press, 3-38.

Darnell, A. C. (1994), A Dictionary of Econometrics. Cheltenham: Edward Elgar.

Fisz, M. (1962), Wahrscheinlichkeitsrechnung und mathematische Statistik. Berlin: Deutscher Verlag der Wissenschaften.

Flaig, G., Licht, G. und V. Steiner (1992), Testing for State Dependence Effects in a Dynamic Model of Male Unemployment Behavior. Zentrum für Europäische Wirtschaftsforschung GmBH, Discussion Paper No. 93-07. Mannheim.

Fletcher, R. (1980), Practical Methods of Optimization. New York: Wiley.

Greene, W. H. (1990), Econometric Analysis. New York: MacMillan.

Hamerle, A. (1994), Panel-Modelle für qualitative Daten. *Allgemeines Statistisches Archiv,* 78, 1-20.

Hamerle, A. und G. Ronning (1994), Panel Analysis for Qualitative Variables. In: Arminger, G., C. C. Clogg and M. E. Sobel, (eds.), Handbook of Statistical Modeling for the Behavioral Sciences. New York: Plenum.

Heckman, J. J. (1981a), Statistical Models for Discrete Panel Data. In: Manski, C. F. and D. Mc Fadden, (eds.), Structural Analysis of Discrete Data with Economic Applications. Cambridge (Mass.): MIT Press, 114-178.

Heckman, J. J. (1981b), The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Stochastic Process. In: Manski, C. F. and D. Mc Fadden, (eds.), Structural Analysis of Discrete Data with Economic Applications. Cambridge (Mass.): MIT Press, 179-195.

Hsiao, C. (1985), Benefits and Limitations of Panel Data. University of Southern California, MRG Working Paper M8505. Los Angeles.

Hsiao, C. (1986), Analysis of Panel Data. Cambridge, Mass: Cambridge University Press.

Hsiao, C. (1992), Logit and Probit Models. In: Matyas, L. and P. Sevestre, (eds.), The Econometrics of Panel Data. Dordrecht: Kluwer, 223-241.

Johnson, N. L. and Kotz, S. (1972), Continuous Multivariate Distributions. New York: Wiley.

Küsters, U. (1987), Hierarchische Mittelwert- und Kovarianzstrukturmodelle mit nichtmetrischen endogenen Variablen. Heidelberg: Physica-Verlag.

Liang, K. Y. and S. L. Zeger (1986), Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73, 13-22.

Maddala, G. S. (1987), Limited Dependent Variable Models Using Panel Data. *Journal of Human Resources*, 12, 307-336.

Rasch, D. (1976), Einführung in die mathematische Statistik. Berlin: Deutscher Verlag der Wissenschaften.

Ronning, G. (1991), Mikroökonometrie. Berlin, Heidelberg: Springer-Verlag.

Schepers, A. und G. Arminger (1992), MECOSA - A Program for the Analysis of General Mean and Covariance Structures with Non-Metric Variables. User Guide. Frauenfeld (CH): SLI-AG.

Schneeweiß, H. (1971), Ökonometrie. Würzburg: Physica-Verlag.

Wagner, G. G., J. Schupp und U. Rendtel (1991), The Socio-Economic Panel (SOEP) for Germany - Methods of Production and Management of Longitudinal Data. Discussion Paper No. 31a. DIW Berlin.

Ziegler, A. und Arminger, G. (1995), Analyzing the Employment Status with Panel Data from the GSOEP. Vierteljahrshefte zur Wirtschaftsforschung, Heft 1, 64. Jahrgang, 72-80.