Institut für Volkswirtschaftslehre und Statistik

No. 588-00

Zur Simulated Maximum-Likelihood-Schätzung von Mehrperioden-Mehralternativen-Probitmodellen

Andreas Ziegler und Angelika Eymann

Beiträge zur angewandten Wirtschaftsforschung

Universität Mannheim
A5, 6
D-68131 Mannheim

Zur Simulated Maximum-Likelihood-Schätzung von Mehrperioden-Mehralternativen-Probitmodellen

Andreas Ziegler und Angelika Eymann*

4. Juli 2000

Zusammenfassung

Die Simulated Maximum-Likelihood-Methode erlaubt die Schätzung von multinomialen Probitmodellen für Querschnitts- oder Paneldaten bei einer größeren Zahl von Entscheidungsalternativen. Im Rahmen von Monte-Carlo-Studien wird hier gezeigt, daß die Simulated Maximum-Likelihood-Methode bei Anwendung des Geweke-Hajivassiliou-Keane-Simulators auch für eine relativ geringe Anzahl an Zufallsziehungen zu exakten Schätzern der Koeffizienten der erklärenden Variablen im Mehrperioden-Mehralternativen-Probitmodell führt. Die Präzision der Schätzung der Varianz-Kovarianz-Parameter scheint dagegen sensitiv bezüglich der Größe der individuellen Auswahlwahrscheinlichkeiten, d.h. bezüglich der Anzahl der Entscheidungsalternativen und der Panelwellen zu sein. In einem einperiodigen Mehralternativen-Probitmodell lassen sich die Varianz-Kovarianz-Parameter relativ exakt bestimmen. Die Schätzgenauigkeit der Varianz-Kovarianz-Parameter im Mehrperioden-Mehralternativen-Probitmodell scheint dagegen mit zunehmender Anzahl der Perioden und Alternativen zu sinken.

1 Einleitung

In den Wirtschafts- und Sozialwissenschaften hat sich das multinomiale Logitmodell als robustes "Arbeitspferd" der empirischen Analyse von Entscheidungen zwischen mehreren wechselseitig ausschließlichen Alternativen seit nunmehr fast drei Jahrzehnten bewährt. Das Modell ist inzwischen in den meisten ökonometrischen Programmpaketen (beispielsweise STATA, SYSTAT, TSP oder LIMDEP) implementiert und gehört zum Standardrepertoire ökonometrischer Lehrbücher.

*Wir danken Axel Börsch-Supan, Elke Eberts und Klaus Winckler für Ihre hilfreichen Kommentare. Diese Forschungsarbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen des Sonderforschungsbereichs 504 gefördert.

Zur Analyse von mehrdimensionalen Entscheidungen, z.B. im Rahmen von Panelmodellen, oder zur Analyse von Entscheidungen zwischen vielen, wechselseitig ausschließlichen Alternativen ist das multinomiale Logitmodell aufgrund seiner restriktiven Annahmen bezüglich der Varianz-Kovarianz-Matrix der stochastischen Modellkomponenten nur bedingt geeignet. In der ökonometrischen, der regionalökonomischen und der Marketing-Literatur wurden diverse Erweiterungen des Logitmodells vorgeschlagen, die eine flexiblere Varianz-Kovarianz-Struktur erlauben, ohne die schätztechnischen Vorzüge des multinomialen Logitmodells ganz aufzugeben. Das genistete Logitmodell (vgl. McFadden, 1978), das auch in einigen Programmpaketen implementiert ist, stellt das bekannteste Beispiel dar. Zahlreiche weitere Erweiterungen, wie beispielsweise "universal logit"-Modelle (vgl. z.B. Fotheringham, 1983, Borgers/Timmermans, 1987, Eymann/Ronning, 1997) oder "random coefficient logit"- bzw. "mixed-logit"-Modelle (vgl. z.B. Guadagni/Little, 1983, Gönül/Srinivasan, 1993, Jain u.a., 1994, Harris/Keane, 1999, McFadden/Train, 2000) haben bisher keinen Eingang in kommerzielle Programmpakete und somit wenig generelle Beachtung gefunden.

Das multinomiale Probitmodell erlaubt eine völlig flexible Struktur der Varianz-Kovarianz-Matrix und ist daher besonders für Panelanalysen, Mehralternativen-Betrachtungen und zur Abbildung unbeobachtbarer individueller Heterogenität geeignet. Nun war die Verwendung derartiger flexibler Probitmodelle bei einer größeren Anzahl an Alternativen und/oder Perioden wegen auftauchender Vielfachintegrale lange Zeit nicht handhabbar. Auch nachdem Simulationsschätzverfahren entwickelt wurden, die die Schätzung dieser Modelle erlauben (vgl. z.B. Lerman/Manski, 1981, McFadden, 1989, Börsch-Supan/Hajivassiliou, 1993, Börsch-Supan, 1994, Keane, 1994), wurden multinomiale Probitmodelle aufgrund des hohen Programmier- und Rechenzeitaufwands für empirische Anwendungen relativ selten genutzt (so z.B. Börsch-Supan u.a., 1992, Börsch-Supan/Pfeiffer, 1992, Chintagunta, 1992, Bolduc u.a., 1996, Alvarez/Nagler, 1997, Asea/Turnovsky, 1998, Bertaut/Starr-McCluer, 2000). Seit kurzem sind Simulationsschätzverfahren für Probitmodelle in einigen Programmpaketen (siehe z.B. GAUSSX und LIMDEP) implementiert, die Eintrittsbarrieren zur empirischen Anwendung von multinomialen Probitmodellen sind daher stark reduziert. Eine genaue Kenntnis der Modellstruktur ist jedoch nach wie vor notwendig, um beispielsweise in LIMDEP die notwendigen Varianz-Kovarianz-Restriktionen vorzugeben.

Der vorliegende Artikel wendet sich vorrangig an Anwenderinnen und Anwender von multinomialen Probitmodellen, die die neu implementierten Routinen in ökonometrischen Programmpaketen nutzen möchten. Im ersten Teil des Papiers werden die ökonometrisch-statistischen Grundlagen zur Schätzung von Mehrperioden-Mehralternativen-Probitmodellen überblicksartig dargestellt. Besondere Aufmerksamkeit wird der Struktur der Varianz-Kovarianz-Matrix der stochastischen Komponenten und den notwendigen, vom Anwender zu definierenden, Restriktionen geschenkt.

Die Darstellung bezieht sich ausschließlich auf die Simulated Maximum-Likelihood-Methode (SMLM) bei Anwendung des Geweke-Hajivassiliou-Keane (GHK)- bzw. Smooth-Recursive-Conditioning (SRC)-Simulators (vgl. Börsch-Supan/Hajivassiliou, 1993 oder Börsch-Supan, 1994). Die methodische Auswahl orientiert sich an den in LIMDEP und GAUSSX implementierten Schätzverfahren und ist letztlich begründet durch die günstigen numerischen Eigenschaften der SMLM sowie durch die hohe Präzision des GHK-Simulators (vgl. Hajivassiliou u.a., 1996).

Im zweiten Teil des Papiers werden Monte-Carlo-Studien vorgestellt, die den Zusammenhang zwischen der Anzahl der Alternativen und Panelwellen und der Präzision der geschätzten Koeffizienten der erklärenden Variablen sowie der geschätzten Varianz-Kovarianz-Parameter im Probitmodell untersuchen. Die vergleichenden Simulationsstudien sollen einen Eindruck vermitteln über die Verläßlichkeit der SMLM-Schätzergebnisse bei Anwendungen mit sehr kleinen individuellen Auswahlwahrscheinlichkeiten, wie sie für Mehrperioden-Mehralternativen-Probitmodelle typisch sind. Sie sollen zudem Hilfestellung geben bei der Wahl der adäquaten (vom Anwender zu definierenden und von der Struktur des Modells abhängigen) Zahl der Zufallsziehungen im GHK-Simulator.

Die Monte-Carlo-Studien orientieren sich an der Vorgehensweise von Lee (1997) für binäre Mehrperioden-Probitmodelle sowie von Geweke u.a. (1997) für Mehrperioden-Mehralternativen-Probitmodelle. Im Gegensatz zu letzteren werden hier SMLM-Schätzergebnisse in verschieden strukturierten ein- und mehrperiodigen Mehralternativen-Probitmodellen gegenübergestellt. Somit kann der kombinierte Effekt einer Zunahme der Anzahl an Perioden und Entscheidungsalternativen im Mehrperioden-Mehralternativen-Probitmodell untersucht werden. Zudem betrachtet die vergleichende Analyse die Auswirkungen unterschiedlicher Beobachtungsumfänge bei konstanter Zahl der Zufallsziehungen auf die Präzision der SMLM-Schätzergebnisse. Die Monte-Carlo-Studien erlauben somit eine Einschätzung des notwendigen Beobachtungsumfangs von Datensätzen, für den die SMLM verläßliche Schätzergebnisse bezüglich der Parameter eines Mehrperioden-Mehralternativen-Probitmodells liefert.

Das Papier ist wie folgt strukturiert: Im zweiten Abschnitt wird zunächst ein Überblick über diskrete Entscheidungsmodelle gegeben und die Struktur des Mehrperioden-Mehralternativen-Probitmodells beschrieben. Im dritten Abschnitt werden Simulationsverfahren zur Approximation von Wahrscheinlichkeiten, die durch Vielfachintegrale gekennzeichnet sind, erläutert. Die Simulated Maximum-Likelihood-Methode zur Schätzung von Mehrperioden-Mehralternativen-Probitmodellen wird anschließend im vierten Abschnitt dargestellt. Die Vorgehensweise der Monte-Carlo-Studien und ihre zentralen Ergebnisse werden im abschließenden fünften Abschnitt diskutiert.

2 Diskrete Entscheidungsmodelle

2.1 Stochastische Nutzenmaximierung

Ausgangspunkt der ökonomischen Ableitung diskreter Mehrperioden-Mehralternativen-Entscheidungsmodelle ist die Möglichkeit einer Untersuchungseinheit i (z.B. einer Person oder eines Unternehmens), zum Zeitpunkt t unter J verschiedenen, sich wechselseitig ausschließenden Alternativen zu wählen. Zum Beispiel kann sich ein Individuum bei der Wohnformwahl in jeder Periode zwischen verschiedenen Wohnformen entscheiden. In dieser Arbeit wird für eine Beobachtungseinheit i zum Zeitpunkt t bzgl. Alternative j folgende hypothetische Nutzenfunktion zugrunde gelegt:

$$\upsilon_{ijt} = \beta_j' x_{it} + \gamma' z_{ijt} + \varepsilon_{ijt} \qquad i = 1, \ldots, N; j = 1, \ldots, J; t = 1, \ldots, T \tag{1}$$

Dabei gliedern sich die erklärenden Variablen in einen K_1-dimensionalen Vektor $x_{it} = (x_{it1}, \ldots, x_{itK_1})'$ von Charakteristika, der über alle Kategorien $j = 1, \ldots, J$ konstant ist, und in einen K_2-dimensionalen Vektor $z_{ijt} = (z_{ijt1}, \ldots, z_{ijtK_2})'$ alternativenspezifischer Attribute bezogen auf die Untersuchungseinheit i. Für die Parametervektoren gilt $\beta_j = (\beta_{j1}, \ldots, \beta_{jK_1})'$ und $\gamma = (\gamma_1, \ldots, \gamma_{K_2})'$. Im folgenden werden die z_{ijt} in den $J \cdot K_2$-dimensionalen Vektor $z_{it} = (z'_{i1t}, \ldots, z'_{iJt})'$ sowie dann die x_{it} und die z_{it} in den $T \cdot (K_1 + J \cdot K_2)$-dimensionalen Vektor $X_i = (x'_{i1}, \ldots, x'_{iT}, z'_{i1}, \ldots, z'_{iT})'$ zusammengefaßt.

2.1.1 Auswahlwahrscheinlichkeiten zum Zeitpunkt t

Der Nutzen υ_{ijt} in (1) kann nicht beobachtet werden und beinhaltet auch eine stochastische Komponente ε_{ijt}, die alle nicht beobachteten Faktoren, welche die Entscheidung für eine Alternative j zum Zeitpunkt t beeinflussen, zusammenfaßt. Beobachtbar sind dagegen die Realisationen folgender Indikatorvariablen ($i = 1, \ldots, N; j = 1, \ldots, J; t = 1, \ldots, T$):

$$D_{ijt} = \begin{cases} 1 & \text{falls Beobachtungseinheit } i \text{ zum Zeitpunkt } t \text{ Kategorie } j \text{ wählt} \\ 0 & \text{sonst} \end{cases}$$

Falls j_{it} die von Untersuchungseinheit i in Periode t tatsächlich gewählte Kategorie darstellt, läßt sich D_{ijt} auch folgendermaßen formulieren ($i = 1, \ldots, N; t = 1, \ldots, T$):

$$D_{ijt} = \begin{cases} 1 & \text{falls } j = j_{it} \in \{1, \ldots, J\} \\ 0 & \text{falls } j \neq j_{it} \end{cases}$$

Gemäß der stochastischen Nutzenmaximierungshypothese (vgl. z.B. Börsch-Supan, 1987 oder Ronning, 1991) entscheidet sich Untersuchungseinheit i zu einem Zeitpunkt t für Kategorie j, falls $\upsilon_{ijt} > \upsilon_{ikt}$ ($\forall k \neq j; k, j = 1, \ldots, J; i = 1, \ldots, N; t = 1, \ldots, T$). Dementsprechend ergeben sich die Auswahlwahrscheinlichkeiten:

$$
\begin{aligned}
P_{ijt} &= P\left(D_{ijt}=1|x_{it},z_{it}\right) \\
&= P\left(\upsilon_{ijt} > \upsilon_{ikt}; \forall k \neq j | x_{it}, z_{it}\right) \\
&= P\left[\varepsilon_{ikt}-\varepsilon_{ijt} < (\beta_j^{'}-\beta_k^{'})x_{it}+\gamma^{'}(z_{ijt}-z_{ikt}); \forall k \neq j\right] \qquad (2)
\end{aligned}
$$

In empirischen Anwendungen wurde in der Vergangenheit häufig der einperiodige Fall ($T = 1$) insbesondere mit $J = 2$ Entscheidungsalternativen betrachtet. Bei derartigen binären Entscheidungsmodellen können die Auswahlwahrscheinlichkeiten mit unterschiedlichen Verteilungsannahmen bzgl. ε_{ij1} relativ einfach dargestellt werden (vgl. z.B. Maddala, 1983, S. 22 ff). Die frühere Konzentration auf einperiodige (binäre oder multinomiale mit $J > 2$ Alternativen) diskrete Entscheidungsmodelle hängt auch damit zusammen, daß lange Zeit kaum entsprechende Paneldaten, bei denen qualitative Variablen über mehrere Perioden beobachtet werden, verfügbar waren.

Das populäre (multinomiale) Logitmodell (vgl. z.B. Ronning, 1991) erhält man aus diesem Modellansatz für $T = 1$ mit der Annahme, daß die stochastischen Komponenten ε_{ij1} über alle Alternativen $j = 1, \ldots, J$ voneinander unabhängig standardextremwertverteilt sind. Die Attraktivität des Logitmodells resultiert aus den einfach und explizit berechenbaren Auswahlwahrscheinlichkeiten P_{ij1}. Das multinomiale Logitmodell besitzt allerdings die Eigenschaft der Unabhängigkeit von irrelevanten Alternativen (IIA-Eigenschaft, vgl. McFadden, 1973 bzw. 1984). Diese Eigenschaft impliziert, daß die Entscheidung zwischen zwei Kategorien unabhängig von der Existenz weiterer Alternativen ist. Diese Annahme ist bei der Analyse vieler Entscheidungssituationen unplausibel. Veranschaulicht wurden die Auswirkungen der IIA-Eigenschaft ursprünglich innerhalb des multiattributiven Logitmodells, das sich durch ausschließlich alternativenspezifische erklärende Variablen auszeichnet. Als Standardbeispiel diente dabei die Wahl zwischen den Verkehrsmitteln Auto, roter Bus und blauer Bus (vgl. z.B. McFadden, 1973, Maddala, 1983, S. 61 f, Maier/Weiss, 1990, S. 141 ff).

2.1.2 Mehrfachintegrale in Auswahlwahrscheinlichkeiten

Die bisherige Spezifikation bezieht sich auf den Spezialfall einer Querschnittsanalyse. Beim Vorliegen von Paneldaten können intertemporale Aspekte in die Betrachtung einbezogen werden (vgl. auch Chamberlain, 1980 bzw. 1984). Eine direkte Übertragung der Annahmen des einperiodigen konventionellen multinomialen Logitmodells auf den mehrperiodigen Fall würde bedeuten, daß die stochastischen Komponenten ε_{ijt} sowohl über alle Kategorien $j = 1, \ldots, J$ als auch über alle Perioden $t = 1, \ldots, T$ unabhängig standardextremwertverteilt sind. Allerdings impliziert dieses Unabhängigkeitspostulat neben der bereits problematisierten IIA-Eigenschaft insbesondere das Fehlen jeglicher intertemporaler Korrelation der stochastischen Nutzenkomponenten ε_{ijt}. Eine solche Annahme ist in vielen Entscheidungssituationen äußerst unrealistisch, vielfach noch unrealistischer als die IIA-Eigenschaft, da sich viele unbeobachtete Einflußfaktoren (zusammengefaßt in der stochastischen Nutzenkomponente ε_{ijt}) im Zeitablauf nur wenig verändern. Häufig hängt die Wahrscheinlichkeit für die Wahl einer bestimmten Alternative j von früheren Erfahrungen ab (vgl. Heckman, 1981). Dementsprechend sollten in der Modellspezifikation beim Vorliegen von Paneldaten intertemporale Abhängigkeiten in den stochastischen Komponenten ε_{ijt} berücksichtigt werden.

Um die rechnerischen Schwierigkeiten eines Entscheidungsmodelles, das beliebige Korrelationen zwischen den stochastischen Komponenten ε_{ijt} gestattet, zu verdeutlichen, soll nun zu den Auswahlwahrscheinlichkeiten P_{ijt} entsprechend (2) zurückgekehrt werden. Ausgehend von bestimmten Verteilungsannahmen können diese allgemein mit der gemeinsamen Dichtefunktion $f_j(\cdot)$ von Differenzen einzelner stochastischer Nutzenkomponenten ε_{ijt} dargestellt werden:

$$P_{ijt} = \int\limits_{-\infty}^{(\beta_j^{'}-\beta_1^{'})x_{it}+\gamma^{'}(z_{ijt}-z_{i1t})} \cdots \int\limits_{-\infty}^{(\beta_j^{'}-\beta_J^{'})x_{it}+\gamma^{'}(z_{ijt}-z_{iJt})} f_j\left(w_{1t},\ldots,w_{j-1,t},w_{j+1,t},\ldots,w_{Jt}\right) dw_{1t}\cdots dw_{j-1,t}\cdot dw_{j+1,t}\cdots dw_{Jt} \qquad (3)$$

Die Wahrscheinlichkeit P_{ijt}, daß Untersuchungseinheit i zu einem bestimmten Zeitpunkt t eine Kategorie j wählt, ist somit in allgemein formulierten einperiodigen Entscheidungsmodellen (ohne vereinfachende Verteilungsannahmen bzgl. der ε_{ijt}) durch $(J-1)$-dimensionale Integrale gekennzeichnet.

Im Zeitablauf steht jede Untersuchungseinheit i vor der Wahl zwischen J^T verschiedenen potentiellen Kategoriensequenzen. Das heißt, im Hinblick auf die tatsächlich gewählte Kategoriensequenz s muß sich eine Beobachtungseinheit i in jeder Periode t (mit $t = 1,\ldots,T$) für eine bestimmte Kategorie j_{it} (mit $j_{it} = 1,\ldots,J$) entscheiden. Die entsprechende Wahrscheinlichkeit P_{is} für die Auswahl einer bestimmten Sequenz s läßt sich auf der Grundlage der stochastischen Nutzenmaximierung in allen Perioden $t = 1,\ldots,T$ folgendermaßen abbilden:

$$\begin{aligned}
P_{is} &= P\left(D_{ij_{i1}1}=1; D_{ij_{i2}2}=1;\ldots;D_{ij_{iT}T}=1|X_i\right)\\
&= P\left(\upsilon_{ij_{it}t} > \upsilon_{ikt}; \forall k \neq j_{it}; k, j_{it}=1,\ldots,J|X_i\right)\\
&= P\left[\varepsilon_{ikt}-\varepsilon_{ij_{it}t} < (\beta_{j_{it}}^{'}-\beta_k^{'})x_{it}+\gamma^{'}(z_{ij_{it}t}-z_{ikt}); \forall k \neq j_{it}; k, j_{it}=1,\ldots,J\right]\\
&= \int\limits_{-\infty}^{(\beta_{j_{i1}}^{'}-\beta_1^{'})x_{i1}+\gamma^{'}(z_{ij_{i1}1}-z_{i11})} \cdots\cdots \int\limits_{-\infty}^{(\beta_{j_{iT}}^{'}-\beta_J^{'})x_{iT}+\gamma^{'}(z_{ij_{iT}T}-z_{iJT})} \qquad (4)\\
&\quad f_s\left(w_{11},\ldots,w_{j_{i1}-1,1},w_{j_{i1}+1,1},\ldots,w_{J1},\ldots\ldots,w_{1T},\ldots,w_{j_{iT}-1,T},w_{j_{iT}+1,T},\ldots,w_{JT}\right)\\
&\quad dw_{11}\cdots dw_{j_{i1}-1,1}\cdot dw_{j_{i1}+1,1}\cdots dw_{J1}\cdots\cdots dw_{1T}\cdots dw_{j_{iT}-1,T}\cdot dw_{j_{iT}+1,T}\cdots dw_{JT}
\end{aligned}$$

Dabei bezeichnet $f_s(\cdot)$ (in Abhängigkeit von Kategoriensequenz s) die gemeinsame Dichtefunktion von Differenzen einzelner stochastischer Nutzenkomponenten ε_{ijt}. Es ist zu erkennen, daß die Wahrscheinlichkeit P_{is} für die Auswahl einer bestimmten Kategoriensequenz s durch Untersuchungseinheit i im allgemeinen durch $(J-1)\cdot T$-dimensionale Integrale gekennzeichnet ist. Sofern keine vereinfachenden Verteilungsannahmen bzgl. der ε_{ijt} getroffen werden (z.B. im mehrperiodigen konventionellen Logitmodell, vgl. oben, aber auch in speziellen Probitmodellen, vgl. nächsten Abschnitt), müssen derartige Vielfachintegrale bei der Schätzung der unbekannten Parameter berechnet werden.

Somit stellt sich die Frage, welche gemeinsame Verteilung den stochastischen Nutzenkomponenten ε_{ijt} auferlegt werden soll. Die Verteilungsannahmen innerhalb des herkömmlichen Logitmodells erfolgten in der Vergangenheit im wesentlichen aufgrund der einfachen Berechnung der Auswahlwahrscheinlichkeiten. In diesem Fall liegen keine Vielfachintegrale vor. Allerdings ist die dem Logitmodell zugrunde liegende Extremwertverteilung auch in verallgemeinerten Ansätzen nicht flexibel genug, um eine beliebige intertemporale und kontemporäre Korrelation zwischen den stochastischen Nutzenkomponenten ε_{ijt} zu gewährleisten. Aus dieser Sicht bietet sich insbesondere die mehrdimensionale Normalverteilung an. Damit gelangt man zu (Mehrperioden-Mehralternativen-)Probitmodellen.

2.2 Mehrperioden-Mehralternativen-Probitmodelle

2.2.1 Varianz-Kovarianz-Restriktionen

Mit der Nutzenfunktion $v_{ijt} = \beta_j' x_{it} + \gamma' z_{ijt} + \varepsilon_{ijt}$ $(i = 1, \ldots, N; j = 1, \ldots, J; t = 1, \ldots, T)$ entsprechend (1) wird im folgenden angenommen:

$$\varepsilon_i = (\varepsilon_{i11}, \ldots, \varepsilon_{iJ1}, \ldots\ldots, \varepsilon_{i1T}, \ldots, \varepsilon_{iJT})' \sim NV(0; \Sigma)$$

Dabei sind die $J \cdot T$-dimensionalen Zufallsvektoren ε_i $(i = 1, \ldots, N)$ untereinander sowie von X_i (d.h. von den erklärenden Variablen) unabhängig.

Unterschiedliche Varianten des Mehrperioden-Mehralternativen-Probitmodells (MMPM) ergeben sich durch verschiedene Restriktionen bzgl. der Varianz-Kovarianz-Matrix Σ. Bei der Parameterschätzung im MMPM sind derartige Restriktionen teilweise zwingend zu beachten. Dies gilt insbesondere bei der Verwendung des Programmpakets LIMDEP, das keine Varianz-Kovarianz-Restriktionen zur Verfügung stellt. Diese müssen vielmehr vom Nutzer selbst einbezogen werden. Zu unterscheiden sind allgemein Restriktionen, die im Hinblick auf die formale Identifikation des MMPM nicht vernachlässigt werden dürfen sowie solche, die lediglich der Reduktion der freien zu schätzenden Parametern dienen. In diesem Abschnitt sollen zunächst die zwingend zu berücksichtigenden Restriktionen diskutiert werden.

Ausgangspunkt sind die $\frac{1}{2} \cdot J \cdot T \cdot (J \cdot T + 1)$ Parameter in Form verschiedener Varianzen und Kovarianzen in Σ. Eine Modellierung, die bei der Parameterschätzung alle diese Varianz-Kovarianz-Parameter (sowie die Parametervektoren β_j, $j = 1, \ldots, J$ und γ) enthält, ist allerdings formal nicht identifiziert (vgl. im folgenden Bolduc, 1992, Bunch, 1991, Dansie, 1985). Dieser Sachverhalt resultiert daraus, daß für die Auswahl einer Kategoriensequenz s durch Beobachtungseinheit i nicht die absoluten Werte der Nutzen v_{ijt} entsprechend (1), sondern vielmehr Nutzendifferenzen entscheidend sind. Damit basiert die Auswahlwahrscheinlichkeit P_{is} nach (4) letztlich nicht auf der Verteilung von ε_i, sondern auf der Verteilung der Differenzen $\varepsilon_{ikt} - \varepsilon_{ij_{it}t}$ $(\forall k \neq j_{it}; k, j_{it} = 1, \ldots, J; t = 1, \ldots, T)$ der stochastischen Nutzenkomponenten. Entsprechend der Verteilungsannahme bzgl. ε_i im MMPM ergibt sich für diese Differenzen:

$$(\varepsilon_{i11} - \varepsilon_{ij_{i1}1}, \ldots, \varepsilon_{iJ1} - \varepsilon_{ij_{i1}1}, \ldots\ldots, \varepsilon_{i1T} - \varepsilon_{ij_{iT}T}, \ldots, \varepsilon_{iJT} - \varepsilon_{ij_{iT}T})' \sim NV(0; \Omega_s)$$

Durch die Unabhängigkeit der ε_i untereinander sind auch diese Vektoren über alle $i = 1, \ldots, N$ unabhängig. Zu erkennen ist die Reduktion der Dimension von Ω_s gegenüber der Dimension von Σ. Das heißt, durch die stochastische Nutzenmaximierung ergibt sich zunächst eine Verringerung der Anzahl der formal identifizierbaren Varianz-Kovarianz-Parameter auf $\frac{1}{2} \cdot [(J-1) \cdot T] \cdot [(J-1) \cdot T + 1]$. Häufige Praxis ist dabei, ausgehend von Σ alle Kovarianzen, die die Kategorie J beinhalten, auf den Wert Null sowie alle Varianzen bzgl. Kategorie J auf den Wert Eins zu restringieren. Darüber hinaus muß (genauso wie z.B. im Logitmodell) wegen der stochastischen Nutzenmaximierung auch ein Parametervektor β_j normiert werden. Anlehnend an die gewählte Basiskategorie J wird im folgenden β_J auf den Nullvektor restringiert.

Hinsichtlich der formalen Identifikation der Varianz-Kovarianz-Parameter ist allerdings noch das Problem der Skalierung zu beachten. Dadurch, daß die Nutzenfunktion v_{ijt} nach (1) lediglich eine latente Variable darstellt, erhält man letztlich durch eine Reskalierung dasselbe Probitmodell. Aus diesem Grund ist eine weitere Fixierung eines Varianz-Kovarianz-Parameters, z.B. einer Varianz, zwingend notwendig. Damit können im allgemeinsten Fall im MMPM maximal $\frac{1}{2} \cdot [(J-1) \cdot T] \cdot [(J-1) \cdot T + 1] - 1$ verschiedene Varianzen und Kovarianzen formal identifiziert werden (vgl. auch Börsch-Supan, 1994) .

Allerdings führt diese allgemeine Spezifikation der Varianz-Kovarianz-Matrix Σ schon bei einer moderaten Anzahl an Kategorien J und/oder Perioden T zu einer sehr hohen Anzahl freier Parameter in Σ. Die Schätzung dieser Vielzahl an Parametern ist in der empirischen Praxis problematisch. Trotz formaler Identifikation lassen sich die Parameter in typischen Beobachtungsumfängen N bei der Schätzung nur schwer identifizieren. Aus diesem Grund werden der Varianz-Kovarianz-Matrix Σ im MMPM häufig Strukturen auferlegt, die sich auf die kontemporäre (vgl. Bolduc, 1992), vor allem aber auf die intertemporale Verknüpfung der unbeobachteten Einflußfaktoren beziehen. Damit kann die Anzahl der freien Varianz-Kovarianz-Parameter reduziert werden. Denkbar wären hier zahlreiche verschiedene Modellierungen. Der folgende sehr allgemeine Ansatz orientiert sich an der gängigen Praxis (vgl. z.B. Börsch-Supan u.a., 1992) und ist insbesondere auch im Programmpaket LIMDEP implementiert.

2.2.2 Eine allgemeine Modellierung

Die stochastische Nutzenkomponente ε_{ijt} gewährleistet hierbei beliebige kontemporäre Verknüpfungen zwischen den Alternativen $j = 1, \ldots, J$. In dieser Hinsicht wird die Anzahl der Varianz-Kovarianz-Parameter zunächst nicht verringert. Die Reduktion der Anzahl dieser Parameter ergibt sich vielmehr aus der zugrunde gelegten intertemporalen Struktur der stochastischen Komponenten ε_{ijt}.

Der hier dargestellte Modellansatz beinhaltet zum einen die bei der Panelanalyse üblicherweise betrachteten stochastischen Effekte. Durch die Berücksichtigung eines solchen zeitinvarianten Zufallseffekts soll die unbeobachtete Heterogenität zwischen einzelnen Untersuchungseinheiten $i = 1, \ldots, N$ erfaßt werden.

Zum anderen wird bei der Modellierung ein autoregressiver Prozeß erster Ordnung einbezogen. Auch die Berücksichtigung dieser Komponente lehnt sich eng an die übliche Modellierung in Regressions- bzw. Zeitreihenmodellen an. Damit kann ein im Zeitablauf systematisch abklingender Einfluß unbeobachteter Faktoren erfaßt werden. Dabei werden im betrachteten MMPM für die verschiedenen Entscheidungsalternativen unterschiedliche autoregressive Verknüpfungen zugelassen. Zu betonen ist, daß die Einbeziehung der autoregressiven Zufallskomponenten sowie der zeitinvarianten stochastischen Effekte in das Probitmodell gegenüber der Einbeziehung im linearen Regressionsmodell eine komplexere Struktur besitzt, da die kontemporären Verknüpfungen zwischen den einzelnen Kategorien berücksichtigt werden müssen.

Konkret gilt entsprechend dieser Ausführungen für die stochastische Nutzenkomponente:

$$\varepsilon_{ijt} = \alpha_{ij} + \nu_{ijt} \qquad i = 1, \ldots, N; j = 1, \ldots, J; t = 1, \ldots, T \tag{5}$$

Die Einbeziehung des autoregressiven Prozesses erster Ordnung erfolgt in der Komponente ν_{ijt}:

$$\nu_{ijt} = \rho_j \nu_{ijt-1} + \sqrt{1 - {\rho_j}^2}\eta_{ijt} \tag{6}$$

Durch das Zurückführen des stochastischen Prozesses auf ν_{ij0} sowie mit der Annahme $\nu_{ij0} = \eta_{ij0}$ erhält man:

$$\nu_{ijt} = \sqrt{1 - {\rho_j}^2} \sum_{m=0}^{t-1} {\rho_j}^m \eta_{ijt-m} + {\rho_j}^t \eta_{ij0} \tag{7}$$

Damit ergibt sich z.B. $\nu_{ij1} = \sqrt{1 - {\rho_j}^2}\, \eta_{ij1} + \rho_j \eta_{ij0}$ oder $\nu_{ij2} = \sqrt{1 - {\rho_j}^2}(\eta_{ij2} + \rho_j \eta_{ij1}) + {\rho_j}^2 \eta_{ij0}$. Mit der Komponente η_{ijt} werden kontemporäre Verknüpfungen zwischen den unbeobachteten Einflußfaktoren berücksichtigt. Es gilt für $t = 0, 1, \ldots, T$:

$$\eta_{ijt} \sim NV(0; \sigma^2_{\eta_j})$$

Dabei sind die η_{ijt} über alle Perioden unkorreliert, d.h. $cov(\eta_{ijt}, \eta_{ij't'}) = 0$ mit $t \neq t'$ und $\forall j, j'$. Für $t = 1, \ldots, T$ gilt $(\forall j, j')$:

$$cov(\eta_{ijt}, \eta_{ij't}) = \sigma_{\eta_{jj'}}$$

Für die ν_{ijt} ist $var(\nu_{ijt}) = \sigma^2_{\nu_j}$ und $cov(\nu_{ijt}, \nu_{ij't}) = \sigma_{\nu_{jj'}}$ $(t = 0, 1, \ldots, T; \forall j, j')$. Darüber hinaus ergibt sich für $t = 1, \ldots, T$ $cov(\eta_{ijt}, \eta_{ij't}) \neq cov(\eta_{ij0}, \eta_{ij'0})$. Aus dem stochastischen Prozeß (6) bzw. (7) folgt insbesondere die gewünschte Annahme $\sigma^2_{\eta_j} = \sigma^2_{\nu_j}$. Mit ρ_j (wobei $|\rho_j| < 1$) wird der Autokorrelationskoeffizient für Kategorie j bezeichnet. Die Einbeziehung zeitinvarianter stochastischer Effekte erfolgt schließlich durch die Komponente α_{ij}. Dabei gilt

$$\alpha_{ij} \sim NV(0; \sigma^2_{\alpha_j})$$

mit

$$cov(\alpha_{ij}, \alpha_{ij'}) = \sigma_{\alpha_{jj'}}$$

wobei die α_{ij} und ν_{ijt} miteinander unkorreliert sind. Letztlich folgt für die Komponenten der Varianz-Kovarianz-Matrix Σ von ε_i (mit $i = 1, \ldots, N; j, j' = 1, \ldots, J; t, t' = 1, \ldots, T$ und $t \geq t'$):

$$cov(\varepsilon_{ijt}, \varepsilon_{ij't'}) = \sigma_{\alpha_{jj'}} + \rho_j{}^{(t-t')} \frac{\sqrt{1-\rho_j{}^2}\sqrt{1-\rho_{j'}{}^2}}{1-\rho_j\rho_{j'}} \sigma_{\eta_{jj'}} \tag{8}$$

Bestandteile dieses zunächst komplex anmutenden Ausdrucks sind ausschließlich die Parameter der oben diskutierten kontemporären, zeitinvarianten sowie autoregressiven Verknüpfungen der unbeobachteten Einflußfaktoren. Für $j = j'$ ergibt sich $(\forall t, t')$ speziell $cov(\varepsilon_{ijt}, \varepsilon_{ijt'}) = \sigma^2_{\alpha_j} + \rho_j{}^{|t-t'|} \sigma^2_{\eta_j}$.

Ein typischer Block der Varianz-Kovarianz-Matrix Σ lautet z.B. für $t = 3$ und $t' = 1$ damit:

$$\begin{matrix}
\sigma^2_{\alpha_1} + \rho_1^2 \sigma^2_{\eta_1} & \sigma_{\alpha_{12}} + \rho_1^2 \frac{\sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}}{1-\rho_1\rho_2} \sigma_{\eta_{12}} & \cdots & \sigma_{\alpha_{1J}} + \rho_1^2 \frac{\sqrt{1-\rho_1^2}\sqrt{1-\rho_J^2}}{1-\rho_1\rho_J} \sigma_{\eta_{1J}} \\
\sigma_{\alpha_{12}} + \rho_2^2 \frac{\sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}}{1-\rho_1\rho_2} \sigma_{\eta_{12}} & \sigma^2_{\alpha_2} + \rho_2^2 \sigma^2_{\eta_2} & \cdots & \sigma_{\alpha_{2J}} + \rho_2^2 \frac{\sqrt{1-\rho_2^2}\sqrt{1-\rho_J^2}}{1-\rho_2\rho_J} \sigma_{\eta_{2J}} \\
\sigma_{\alpha_{13}} + \rho_3^2 \frac{\sqrt{1-\rho_1^2}\sqrt{1-\rho_3^2}}{1-\rho_1\rho_3} \sigma_{\eta_{13}} & \sigma_{\alpha_{23}} + \rho_3^2 \frac{\sqrt{1-\rho_2^2}\sqrt{1-\rho_3^2}}{1-\rho_2\rho_3} \sigma_{\eta_{23}} & \cdots & \sigma_{\alpha_{3J}} + \rho_3^2 \frac{\sqrt{1-\rho_3^2}\sqrt{1-\rho_J^2}}{1-\rho_3\rho_J} \sigma_{\eta_{3J}} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_{\alpha_{1J}} + \rho_J^2 \frac{\sqrt{1-\rho_1^2}\sqrt{1-\rho_J^2}}{1-\rho_1\rho_J} \sigma_{\eta_{1J}} & \sigma_{\alpha_{2J}} + \rho_J^2 \frac{\sqrt{1-\rho_2^2}\sqrt{1-\rho_J^2}}{1-\rho_2\rho_J} \sigma_{\eta_{2J}} & \cdots & \sigma^2_{\alpha_J} + \rho_J^2 \sigma^2_{\eta_J}
\end{matrix}$$

Zu beachten ist, daß bei dieser Modellierung entsprechend der Ausführungen im vorhergehenden Abschnitt hinsichtlich der formalen Identifikation eine Reihe von Restriktionen vorgenommen werden müssen. Aus der stochastischen Nutzenmaximierung ergibt sich, daß $J-1$ Varianzparameter und $\frac{1}{2} \cdot (J-1) \cdot (J-2)$ Kovarianzparameter der kontemporären Verknüpfungen, $J-1$ Varianzparameter und keine Kovarianzparameter der zeitinvarianten stochastischen Effekte (α_{ij}) sowie $J-1$ Autokorrelationsparameter (ρ_j) formal identifizierbar sind. Darüber hinaus muß hinsichtlich des Skalierungsproblems ein weiterer Varianzparameter der kontemporären Verknüpfungen restringiert werden, so daß tatsächlich lediglich $J-2$ derartige Varianzparameter (σ_{η_j}) formal zu identifizieren sind. Damit ist die Varianz-Kovarianz-Matrix Σ durch insgesamt $\frac{1}{2} \cdot (J^2 + 3 \cdot J - 6)$ freie (zu schätzende) Parameter gekennzeichnet.

Falls man erneut J als Basiskategorie betrachtet, können $\sigma^2_{\eta_J}$, $\sigma^2_{\eta_{J-1}}$ und $\sigma^2_{\alpha_J}$ auf den Wert Eins sowie $\sigma_{\eta_{jJ}}$ $(\forall j \neq J)$, $\sigma_{\alpha_{jj'}}$ $(\forall j \neq j')$ und ρ_J auf den Wert Null normiert werden (vgl. auch Börsch-Supan u.a., 1992). Bei dieser Form der Restriktion ist zu beachten, daß dem MMPM eine zeitinvariante Verknüpfung auferlegt wird. Das heißt, durch keinerlei Struktur der freien Parameter kann diese Form der Korrelation beseitigt werden. Dieses Problem wird umgangen, indem $\sigma^2_{\alpha_J}$ auf den Wert Null restringiert wird. Durch die Normierung hinsichtlich ρ_J erscheinen bei der intertemporalen autoregressiven Korrelation keine derartige Schwierigkeiten.

Zu erwähnen ist, daß in die Parameterschätzung häufig (vgl. auch die Monte-Carlo-Studien im fünften Abschnitt) anstatt der Varianzen $\sigma^2_{\eta_j}$ $(j = 1, \ldots, J-2)$ und $\sigma^2_{\alpha_j}$ $(j = 1, \ldots, J-1)$ die entsprechenden Standardabweichungen σ_{η_j} bzw. σ_{α_j} und anstatt der Kovarianzen $\sigma_{\eta_{jj'}}$ $(j, j' = 1, \ldots, J-1; j \neq j')$ die Korrelationskoeffizienten $corr(\eta_{ijt}, \eta_{ij't}) = \sigma_{\eta_{jj'}}/\sigma_{\eta_j}\sigma_{\eta_{j'}}$ eingehen.

Diese beschriebene allgemeine Spezifikation beinhaltet viele einfache Probitmodelle als Spezialfälle. Zum Beispiel werden mit $\alpha_i = (\alpha_{i1}, \ldots, \alpha_{iJ})' = 0$ die zeitinvarianten stochastischen Effekte und mit $\rho = (\rho_1, \ldots, \rho_J)' = 0$ die autoregressiven intertemporalen Verknüpfungen beseitigt. Falls $\alpha_i = 0$ und $\rho = 0$, bildet das MMPM keinerlei intertemporale Korrelationen mehr ab. Mit $T = 1$ (sowie $\alpha_i = 0$ und $\rho = 0$) gelangt man demnach zu einem einperiodigen Mehralternativen-Probitmodell, das (im allgemeinen) eine kontemporäre Verknüpfung der ε_{ij1} über alle Kategorien $j = 1, \ldots, J$ zuläßt. Falls $T = 1$ (sowie $\alpha_i = 0$ und $\rho = 0$) und die Varianz-Kovarianz-Matrix von $\eta_{i1} = (\eta_{i11}, \ldots, \eta_{iJ1})'$ die Einheitsmatrix I darstellt, erhält man schließlich speziell das einperiodige Independent Probitmodell.

Auf der Grundlage des diskutierten allgemein formulierten MMPM werden in den Monte-Carlo-Studien im fünften Abschnitt im Hinblick auf die Untersuchung praktischer Aspekte der Simulated Maximum-Likelihood-Schätzung ein- bzw. mehrperiodige Mehralternativen-Probitmodelle analysiert. Im folgenden werden alle freien Parameter des jeweils betrachteten

Probitmodells im Vektor θ zusammengefaßt. Der wahre, unbekannte und zu schätzende Parametervektor wird speziell mit $\dot{\theta}$ bezeichnet.

3 Simulationsverfahren

Im letzten Abschnitt wurde gezeigt, daß die (vom Parametervektor θ abhängigen) Auswahlwahrscheinlichkeiten $P_{ijt}(\theta)$ in (3) sowie insbesondere $P_{is}(\theta)$ in (4) häufig durch Vielfachintegrale gekennzeichnet sind. Bei der Parameterschätzung wird sich zeigen, daß die wiederholte Berechnung derartiger Auswahlwahrscheinlichkeiten notwendig ist. Die Approximation dieser Wahrscheinlichkeiten mit herkömmlichen (deterministischen) numerischen Integrationsmethoden, die keine Zufallselemente einschließen, ist jedoch oft bei flexiblen Probitmodellen rechnerisch nicht möglich. Stattdessen können die Wahrscheinlichkeiten schnell und genau mit Hilfe von (stochastischen) Simulationsmethoden, d.h. mit transformierten Ziehungen von Pseudo-Zufallszahlen, geschätzt werden (vgl. z.B. Hajivassiliou u.a., 1996). Allgemeiner Ausgangspunkt der rechnerischen Probleme bei der Parameterschätzung im allgemein formulierten MMPM ist ein Zufallsvektor $U = (U_1, \ldots, U_H)'$ mit $U \sim NV(0;\Omega)$. Im Hinblick auf die Auswahl einer bestimmten Kategoriensequenz s durch Untersuchungseinheit i gilt speziell $H = (J-1) \cdot T$. Zu berechnen sind also mit den Konstanten a_h und b_h (wobei $a_h < b_h; h = 1, \ldots, H$) folgende typische Wahrscheinlichkeiten:

$$P\left(a_1 < U_1 < b_1; \ldots; a_H < U_H < b_H\right) = \int_{a_H}^{b_H} \cdots \int_{a_1}^{b_1} f\left(u_1, \ldots, u_H\right) du_1 \cdots du_H \tag{9}$$

Dabei bezeichnet $f(u) = f(u_1, \ldots, u_H)$ die Dichtefunktion des normalverteilten Zufallsvektors U. Zur Approximation dieser analytisch nicht handhabbaren Wahrscheinlichkeit wird zunächst eine ganze Klasse von Simulationsverfahren vorgeschlagen.

3.1 Importance-Sampling

Das Prinzip dieser Simulationstechnik wird durch die modifizierte Darstellung der Wahrscheinlichkeit in (9) deutlich (vgl. Hajivassiliou u.a., 1996):

$$\begin{aligned}
P(a_1 < U_1 < b_1; \ldots; a_H < U_H < b_H) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{1}(u \in A) f(u) du \\
&= E\left[\mathbf{1}(U \in A)\right] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left[\mathbf{1}(v \in A) \frac{f(v)}{g(v)}\right] g(v) dv \\
&= E\left[\mathbf{1}(V \in A) \frac{f(V)}{g(V)}\right]
\end{aligned}$$

Dabei stellt $\mathbf{1}(U \in A)$ die Indikatorfunktion für das Ereignis $A = \{a < U < b\}$ mit $a = (a_1, \ldots, a_H)'$ und $b = (b_1, \ldots, b_H)'$. Es bezeichnet $g(v)$ die Dichtefunktion eines Zufallsvektors $V = (V_1, \ldots, V_H)'$ mit einem positiven Träger in einer Menge, die A umfaßt (vgl. auch Börsch-Supan, 1994). Mit der Bestimmung der relativen Häufigkeit der Ereignisse $U_r \in A$, wobei die Zufallsvektoren U_r $(r = 1, \ldots, R)$ mit Hilfe von Zufallszahlengeneratoren aus der Verteilung von U gezogen werden, gelangt man zum einfachen Häufigkeitssimulator. Es hat sich jedoch in vergleichenden Monte-Carlo-Studien (z.B. Börsch-Supan/Hajivassiliou, 1993, Hajivassiliou u.a., 1996) herausgestellt, daß dieser Häufigkeitssimulator vielen anderen Simulationsmethoden hinsichtlich der Annäherung an die wahre Wahrscheinlichkeit unterlegen ist. Mit den aus der Verteilung von V gezogenen Zufallsvektoren $V_r = (V_{1r}, \ldots, V_{Hr})'$ $(r = 1, \ldots, R)$ gelangt man dagegen zu einem Importance-Sampling-Simulator:

$$\widetilde{P}^{IS}(a_1 < U_1 < b_1; \ldots; a_H < U_H < b_H) = \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}(V_r \in A)\frac{f(V_r)}{g(V_r)} \tag{10}$$

Dabei bezeichnet R die Anzahl der jeweiligen Zufallsziehungen bzw. Simulationsreplikationen. Zu erkennen ist, daß der Importance-Sampling-Simulator nicht auf Ziehungen aus der Verteilung des zugrunde liegenden (normalverteilten) Zufallsvektors U basiert (vgl. z.B. Hajivassiliou u.a., 1996). Die Importance-Dichtefunktion $g(v)$ sollte derart gewählt werden, daß $P(V_r \in A)$ möglichst groß ist. Im idealen Fall besitzt $g(v)$ genau den Träger A, so daß $P(V_r \in A) = 1$, wodurch in (10) die Indikatorfunktion $\mathbf{1}(\cdot)$ vernachlässigt werden kann. Eine wichtige Rolle spielt auch, daß sich die Dichtefunktion $g(v)$ im Intervall A der Dichtefunktion $f(u)$ gut anpaßt (vgl. z.B. Kaltenborn, 1997, S. 63 ff).

Die Simulation der Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ aus (4) lautet demnach mit den Replikationen $V_{ir}^s = (V_{i1r}^s, \ldots, V_{iHr}^s)'$ $(r = 1, \ldots, R)$ der Zufallsvektoren $V_i^s = (V_{i1}^s, \ldots, V_{iH}^s)'$ mit entsprechender Importance-Dichtefunktion $g(v_i^s)$, wobei $H = (J-1) \cdot T$:

$$\begin{aligned} \widetilde{P}_{is}^{IS}(\theta) \;=\; & \frac{1}{R}\sum_{r=1}^{R} \mathbf{1}\Big[-\infty < V_{i1r}^s < (\beta_{j_{i1}}^{'} - \beta_1^{'})x_{i1} + \gamma^{'}(z_{ij_{i1}1} - z_{i11}); \ldots; \\ & -\infty < V_{iHr}^s < (\beta_{j_{iT}}^{'} - \beta_J^{'})x_{iT} + \gamma^{'}(z_{ij_{iT}T} - z_{iJT})\Big] \frac{f(V_{i1r}^s, \ldots, V_{iHr}^s)}{g(V_{i1r}^s, \ldots, V_{iHr}^s)} \end{aligned} \tag{11}$$

Der Importance-Sampling-Simulator ist hinsichtlich $P_{is}(\theta)$ erwartungstreu (vgl. z.B. Wilde, 1999, S. 112). Dabei hängt die (bedingte) Erwartungstreue im MMPM von den erklärenden Variablen in X_i ab. Importance-Sampling-Simulatoren unterscheiden sich (im MMPM) nach der Wahl der Importance-Dichtefunktionen $g(v_i^s)$. Hajivassiliou u.a. (1996) analysieren Importance-Dichtefunktionen, die sich auf unabhängige exponentialverteilte bzw. auf gestutzte normalverteilte Zufallsvariablen beziehen. Vijverberg (1997) extrahiert im Rahmen der Importance-Sampling-Technik eine Klasse rekursiver Simulatoren. Der wichtigste Vertreter dieser Familie ist der Smooth-Recursive-Conditioning (SRC)- bzw. Geweke-Hajivassiliou-Keane (GHK)-Simulator.

3.2 GHK-Simulator

Der GHK-Simulator beruht primär auf der Darstellung der Auswahlwahrscheinlichkeiten als Produkt bedingter Wahrscheinlichkeiten (vgl. z.B. Börsch-Supan, 1994). Darauf aufbauend erfolgt die Ziehung der Vektoren $V_{ir}^s = (V_{i1r}^s, \ldots, V_{iH-1r}^s)'$ $(r = 1, \ldots, R)$ aus der Verteilung von V_i^s mit entsprechender Importance-Dichtefunktion $g(v_i^s)$ durch die sequentielle Ziehung der einzelnen Komponenten V_{ihr}^s $(h = 1, \ldots, H-1)$ von V_{ir}^s aus der gestutzten Standardnormalverteilung. Eine ausführliche Darstellung dieser speziellen Dichtefunktion $g(v_i^s)$ im Rahmen der Importance-Sampling-Technik findet sich in Wilde (1999). Somit können die Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ in (4) mit dem GHK-Simulator folgendermaßen approximiert werden:

$$\tilde{P}_{is}^{GHK}(\theta) = \frac{1}{R}\sum_{r=1}^{R} \Phi\left(\frac{(\beta_{j_{i1}}' - \beta_1')x_{i1} + \gamma'(z_{ij_{i1}1} - z_{i11})}{l_{11}^s}\right) \cdot \tag{12}$$

$$\Phi\left(\frac{(\beta_{j_{i1}}' - \beta_2')x_{i1} + \gamma'(z_{ij_{i1}1} - z_{i21}) - l_{21}^s V_{i1r}^s}{l_{22}^s}\right) \cdots$$

$$\Phi\left(\frac{(\beta_{j_{iT}}' - \beta_J')x_{iT} + \gamma'(z_{ij_{iT}T} - z_{iJT}) - l_{H1}^s V_{i1r}^s - \cdots - l_{HH-1}^s V_{iH-1r}^s}{l_{HH}^s}\right)$$

Dabei bezeichnen l_{km}^s $(k, m = 1, \ldots, H)$ die Elemente der unteren Dreiecksmatrix, die sich aus der Cholesky-Zerlegung der Varianz-Kovarianz-Matrix Ω_s der gemeinsam normalverteilten Differenzen $(\varepsilon_{i11} - \varepsilon_{ij_{i1}1}, \ldots, \ldots, \varepsilon_{iJT} - \varepsilon_{ij_{iT}T})'$ ergibt (vgl. Abschnitt 2.2.1). Da der GHK-Simulator einen speziellen Importance-Sampling-Simulator darstellt, ist auch er hinsichtlich $P_{is}(\theta)$ erwartungstreu. Im Unterschied z.B. zum Häufigkeitssimulator ist der GHK-Simulator stetig und differenzierbar in den Parametern der Wahrscheinlichkeit $P_{is}(\theta)$ für die Wahl einer Kategoriensequenz s im MMPM. Zudem kann der GHK-Simulator nicht den Wert Null annehmen.

In vielen vergleichenden Monte-Carlo-Studien hat sich die Überlegenheit des GHK-Simulators gegenüber anderen Simulationsmethoden hinsichtlich der Approximationsgüte von Wahrscheinlichkeiten herauskristallisiert (vgl. z.B. Börsch-Supan/Hajivassiliou, 1993, Mühleisen, 1994, Hajivassiliou u.a., 1996, Vijverberg, 1997). So können auch sehr kleine Wahrscheinlichkeiten mit dem GHK-Simulator präzise geschätzt werden. Der GHK-Simulator besitzt zudem geringere durchschnittliche Verzerrungen bei der Annäherung an die wahre Wahrscheinlichkeit sowie einen geringeren mittleren quadratischen Fehler als andere Simulationsverfahren. Schließlich genügt schon eine kleine Anzahl R an Simulationsreplikationen, um mit dem GHK-Simulator Wahrscheinlichkeiten sehr genau zu schätzen.

4 Simulated Maximum-Likelihood-Schätzung

Ausgangspunkt der Parameterschätzung in dieser Arbeit sind N unabhängige Beobachtungspaare (Y_i, X_i). Im MMPM sind im Vektor X_i alle erklärenden Variablen (vgl. Abschnitt 2.1) und im J^T-dimensionalen Vektor $Y_i = (Y_{i1}, Y_{i2}, \ldots)'$ sind die beobachtbaren endogenen Variablen enthalten mit:

$$Y_{is} = \begin{cases} 1 & \text{falls Untersuchungseinheit } i \text{ die Kategoriensequenz } s \text{ wählt} \\ 0 & \text{sonst} \end{cases}$$

Dabei gilt $s \in S$, wobei S die Menge aller J^T potentiellen Kategoriensequenzen bezeichnet. Darüber hinaus ist im MMPM Y_i (gegeben X_i) multinomialverteilt mit den Parametern 1 und allen $P_{is}(\theta)$ $(s \in S)$. Somit ergibt sich für den Schätzer entsprechend der Maximum-Likelihood-Methode (MLM):

$$\hat{\theta}_{MLM} = \arg\max_{\theta} \left[lnL(\theta)\right] = \arg\max_{\theta} \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} ln P_{is}(\theta)\right] \tag{13}$$

Damit erfordert die MLM-Schätzung im allgemein formulierten MMPM für jede Beobachtungseinheit i und für jede Iteration im Maximierungsprozeß die Berechnung eines $(J-1) \cdot T$-dimensionalen Integrals zur Bestimmung der Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ entsprechend (4). Eine solche Berechnung ist bei großen J und/oder T (vgl. auch Stern, 1999) mit herkömmlichen numerischen Integrationsverfahren wegen der hohen rechnerischen Kosten nicht handhabbar. Jedoch können die $P_{is}(\theta)$ z.B. durch die Simulatoren in (11) oder (12) approximiert werden. Bei der Verknüpfung des MLM-Ansatzes in (13) mit einem derartigen Simulator $\widetilde{P}_{is}(\theta)$ gelangt man zur Simulated Maximum-Likelihood-Methode (SMLM) (vgl. z.B. Gourieroux/Monfort, 1996) mit dem Schätzer im MMPM:

$$\hat{\theta}_{SMLM} = \arg\max_{\theta} \left\{\sum_{i=1}^{N} \sum_{s \in S} Y_{is} ln \widetilde{P}_{is}(\theta)\right\} \tag{14}$$

Im Hinblick auf die numerische Stabilität ist zu beachten, daß die Verknüpfung der MLM mit einem diskreten Simulator wie z.B. dem Häufigkeitssimulator (bei endlichem R) eine vom Parametervektor θ abhängige diskrete Zielfunktion verursacht. So gelangt man schon bei einer kleinen Modifikation der Parameterwerte zu einer Veränderung der Zielfunktion in diskreten Schritten. Dies kann zu numerischen Problemen bei der Maximierung nach θ führen (vgl. Börsch-Supan/Hajivassiliou, 1993). Zudem sollte man beachten, daß die ganz zu Beginn erzeugten Pseudo-Zufallszahlen (Ausgangspunkt sind häufig Realisationen [0;1]-rechteckverteilter Pseudo-Zufallsvariablen) im iterativen Optimierungsprozeß nicht mit den Parametern verändert werden (vgl. Hajivassiliou/Ruud, 1994 bzw. Stern, 1999).

Alternativ könnte auch eine "method of simulated moments" sowie eine "method of simulated scores" zur Parameterschätzung im MMPM herangezogen werden (zu einer Diskussion der Beziehungen zwischen diesen klassischen Simulationsschätzverfahren speziell in Probitmodellen sowie der Vor- und Nachteile dieser Schätzer vgl. Ziegler, 2000). Im folgenden wird jedoch ausschließlich die SMLM in Verbindung mit dem GHK-Simulator betrachtet, wie sie auch in den beiden Programmpaketen LIMDEP und GAUSSX implementiert ist.

Zur Untersuchung des Verhaltens von SMLM/GHK-Schätzern in Probitmodellen existiert in der Literatur bereits eine Vielzahl an vergleichenden Simulationsstudien. Bisherige Monte-Carlo-Studien betrachten jedoch häufig spezielle Probitmodelle, d.h. binäre mehrperiodige Probitmodelle (vgl. z.B. Keane, 1994, Lee, 1995, 1997, Hyslop, 1999, Inkmann, 1999) oder einperiodige Mehralternativen-Probitmodelle (vgl. z.B. Börsch-Supan/Hajivassiliou, 1993, Geweke u.a., 1994, Stern, 1999). Die einzige (uns bekannte) systematische Untersuchung eines allgemein formulierten MMPM erfolgt in Geweke u.a. (1997). Dabei wird allerdings der für empirische Anwendungen wichtige Aspekt der Auswirkungen des Beobachtungsumfangs N auf SMLM-Schätzungen nicht analysiert. Zudem sind die Aussagen hinsichtlich der Wahl der adäquaten Anzahl R an Simulationsreplikationen in der Literatur recht widersprüchlich.

Die Monte-Carlo-Studien im nächsten Abschnitt versuchen zum einen, potentiellen Anwendern weitere praktische Hinweise für den Umgang mit der SMLM in Probitmodellen, insbesondere hinsichtlich des Beobachtungsumfangs N sowie der Auswahl der Anzahl R an Simulationsreplikationen, zu geben. Dabei wird die SMLM-Schätzung der Parameter der erklärenden Variablen derjenigen der Varianz-Kovarianz-Parameter gegenübergestellt. Zudem werden im folgenden auch erstmals SMLM-Schätzungen in verschieden strukturierten Mehralternativen-Probitmodellen analysiert. Damit wird der Zusammenhang zwischen der Anzahl an Alternativen sowie Panelwellen und der Präzision der geschätzten Parameter im Probitmodell untersucht. Dabei zeigen sich bei den betrachteten einperiodigen Vieralternativen-, fünfperiodigen Dreialternativen- sowie achtperiodigen Vieralternativen-Probitmodellen vor allem hinsichtlich der SMLM-Schätzung der Varianz-Kovarianz-Parameter markante Unterschiede.

5 Monte-Carlo-Studien

Anknüpfend an obige Bemerkungen wurden in den folgenden Untersuchungen die unbekannten Parameter ausschließlich mit der SMLM in Verbindung mit dem GHK-Simulator geschätzt. In den einzelnen 20 bzw. 200 Replikationen des datengenerierenden Prozesses (DGP) eines der verschiedenen Experimente wurden durchweg (auch bei unterschiedlicher Anzahl R an Simulationsreplikationen) dieselben (pseudo-zufällig erzeugten) erklärenden Variablen verwendet. Bei einer Erhöhung des Beobachtungsumfanges N wurden die bei kleinerem N zunächst generierten erklärenden Variablen in die Parameterschätzung aufgenommen. Die (Pseudo-)Zufallszahlen zur stochastischen Integration wurden dagegen über die jeweiligen 20 bzw. 200 Replikationen des DGP modifiziert. Bei einer sukzessiven Erhöhung des

Beobachtungsumfanges N bzw. der Anzahl R an Simulationsreplikationen wurden jedoch die bei kleinerem N bzw. kleinerem R generierten Zufallszahlen entsprechend einbezogen.

5.1 Design

5.1.1 Erstes Experiment: Einperiodiges Vieralternativen-Probitmodell

Zunächst wird ausgehend von (1) folgendes Probitmodell betrachtet ($i = 1, \ldots, N$):

$$\begin{aligned} \upsilon_{i11} &= \gamma_1 z_{i111} + \gamma_2 z_{i112} + \varepsilon_{i11} \\ \upsilon_{i21} &= \gamma_1 z_{i211} + \gamma_2 z_{i212} + \varepsilon_{i21} \\ \upsilon_{i31} &= \gamma_1 z_{i311} + \gamma_2 z_{i312} + \varepsilon_{i31} \\ \upsilon_{i41} &= \gamma_1 z_{i411} + \gamma_2 z_{i412} + \varepsilon_{i41} \end{aligned}$$

Die zwei alternativenspezifischen Attribute wurden wie folgt generiert ($i = 1, \ldots, N; j = 1, \ldots, 4$):

$$z_{ij11} \sim NV(0;2) \quad z_{ij21} \sim NV(0;2)$$

Dabei liegen den entsprechenden Parametern folgende Werte zugrunde:

$$\dot{\gamma}_1 = 1 \quad \dot{\gamma}_2 = 0$$

Hinsichtlich der Varianz-Kovarianz-Komponenten entsprechend (8) beinhaltet der kontemporäre Korrelationen einschließende DGP folgende Parameter (da $T = 1$ gilt $\alpha_i = \rho = 0$):

$$\dot{\sigma}_{\eta_1} = 1.5 \quad \dot{\sigma}_{\eta_2} = 0.5$$

$$\dot{corr}(\eta_{i11}, \eta_{i21}) = \dot{corr}(\eta_{i11}, \eta_{i31}) = \dot{corr}(\eta_{i21}, \eta_{i31}) = 0.5$$

Zur formalen Identifizierung wurde entsprechend Abschnitt 2.2.2 $\dot{\sigma}_{\eta_3}$ und $\dot{\sigma}_{\eta_4}$ auf den Wert Eins sowie $\dot{\sigma}_{\eta_{j4}}$ ($j = 1, 2, 3$) auf den Wert Null restringiert. Der Beobachtungsumfang wurde zwischen $N = 1000$ und $N = 2000$, die Anzahl der Simulationsreplikationen wurde zwischen $R = 10$, $R = 50$ und $R = 200$ variiert. Bei einem geringeren Beobachtungsumfang ($N = 500$) ergaben sich wiederholt Schwierigkeiten bei der Konvergenz der simulierten Loglikelihoodfunktion zu einem Maximum. Die Anzahl der Replikationen des DGP beträgt bei diesem Experiment durchweg 200.

5.1.2 Zweites Experiment: Fünfperiodiges Dreialternativen-Probitmodell

Hinsichtlich der Analyse von Paneldaten wird zunächst folgendes Probitmodell betrachtet ($i = 1, \ldots, N;\ t = 1, \ldots, 5$):

$$\begin{aligned} \upsilon_{i1t} &= \gamma_1 z_{i1t1} + \gamma_2 z_{i1t2} + \varepsilon_{i1t} \\ \upsilon_{i2t} &= \gamma_1 z_{i2t1} + \gamma_2 z_{i2t2} + \varepsilon_{i2t} \\ \upsilon_{i3t} &= \gamma_1 z_{i3t1} + \gamma_2 z_{i3t2} + \varepsilon_{i3t} \end{aligned}$$

Die zwei alternativenspezifischen Attribute wurden im Hinblick auf intertemporale Verknüpfungen wie folgt generiert ($i = 1, \ldots, N; j = 1, \ldots, 3; t = 1, \ldots, 5$):

$$z_{ijt1} = z_{ij1}^{(1)} + z_{ijt1}^{(2)} \quad \text{wobei} \quad z_{ij1}^{(1)} \sim NV(0;1) \quad \text{und} \quad z_{ijt1}^{(2)} \sim NV(0;1)$$

$$z_{ijt2} = z_{ij2}^{(1)} + z_{ijt2}^{(2)} \quad \text{wobei} \quad z_{ij2}^{(1)} \sim NV(0;1) \quad \text{und} \quad z_{ijt2}^{(2)} \sim NV(0;1)$$

Den entsprechenden Parametern liegen folgende Werte zugrunde:

$$\dot{\gamma}_1 = 1 \quad \dot{\gamma}_2 = 0$$

Die Varianz-Kovarianz-Parameter des DGP in diesem allgemein formulierten MMPM lauten:

$$\dot{\sigma}_{\eta_1} = 1.5 \quad co\dot{r}r(\eta_{i1t}, \eta_{i2t}) = 0.5$$

$$\dot{\sigma}_{\alpha_1} = 1.5 \quad \dot{\sigma}_{\alpha_2} = 0.5$$

$$\dot{\rho}_1 = 0.8 \quad \dot{\rho}_2 = 0.5$$

Zur formalen Identifizierung der Parameter wurden die in Abschnitt 2.2.2 zum Ende vorgeschlagenen Normierungen gewählt, d.h. speziell σ_{α_3} wurde auf den Wert Null restringiert. Der Beobachtungsumfang wurde zwischen $N = 250$ und $N = 500$, die Anzahl an Simulationsreplikationen wurde zwischen $R = 10$, $R = 50$ und $R = 200$ variiert. Die Anzahl der Replikationen des DGP beträgt auch bei diesem Experiment jeweils 200.

5.1.3 Drittes Experiment: Achtperiodiges Vieralternativen-Probitmodell

Schließlich wird folgendes Probitmodell betrachtet ($i = 1, \ldots, N$; $t = 1, \ldots, 8$):

$$\begin{aligned}
\upsilon_{i1t} &= \gamma_1 z_{i1t1} + \gamma_2 z_{i1t2} + \varepsilon_{i1t} \\
\upsilon_{i2t} &= \gamma_1 z_{i2t1} + \gamma_2 z_{i2t2} + \varepsilon_{i2t} \\
\upsilon_{i3t} &= \gamma_1 z_{i3t1} + \gamma_2 z_{i3t2} + \varepsilon_{i3t} \\
\upsilon_{i4t} &= \gamma_1 z_{i4t1} + \gamma_2 z_{i4t2} + \varepsilon_{i4t}
\end{aligned}$$

Hinsichtlich der beiden alternativenspezifischen Attribute wurde hier auch eine Dummy-Variable in die Betrachtung einbezogen, d.h. ($i = 1, \ldots, N; j = 1, \ldots, 4; t = 1, \ldots, 8$):

$$z_{ijt1} = z_{ij1}^{(1)} + z_{ijt1}^{(2)} \quad \text{wobei} \quad z_{ij1}^{(1)} \sim NV(0;1) \quad \text{und} \quad z_{ijt1}^{(2)} \sim NV(0;1)$$

$$z_{ijt2} \sim BV(0.5)$$

Den entsprechenden Parametern liegen folgende Werte zugrunde:

$$\dot{\gamma}_1 = 1 \quad \dot{\gamma}_2 = 1$$

Die Varianz-Kovarianz-Parameter des DGP in diesem allgemein formulierten MMPM lauten:

$$\dot{\sigma}_{\eta_1} = 1.5 \quad \dot{\sigma}_{\eta_2} = 0.5$$

$$\dot{corr}(\eta_{i1t}, \eta_{i2t}) = \dot{corr}(\eta_{i1t}, \eta_{i3t}) = \dot{corr}(\eta_{i2t}, \eta_{i3t}) = 0.5$$

$$\dot{\sigma}_{\alpha_1} = 1.5 \quad \dot{\sigma}_{\alpha_2} = 1.5 \quad \dot{\sigma}_{\alpha_3} = 0.5$$

$$\dot{\rho}_1 = 0.8 \quad \dot{\rho}_2 = 0.8 \quad \dot{\rho}_3 = 0.5$$

Zur formalen Identifizierung der Parameter wurden die in Abschnitt 2.2.2 (vgl. auch Börsch-Supan u.a., 1992) zuerst vorgeschlagenen Normierungen gewählt, d.h. σ_{α_4} wurde auf den Wert Eins restringiert. Der Beobachtungsumfang wurde zwischen $N = 250$ und $N = 500$, die Anzahl der Simulationsreplikationen wurde zwischen $R = 10$, $R = 50$ und $R = 200$ variiert. Die Anzahl der Replikationen des DGP beträgt bei diesem Experiment aufgrund der langen Rechenzeiten jeweils 20.

5.2 Ergebnisse

Im folgenden werden mit $\overline{|Bias|}$ der durchschnittliche Betrag der Verzerrungen und mit $\overline{Rmse}$ der Durchschnitt der Wurzel aus der mittleren quadratischen Abweichung aller Parameterschätzwerte bezogen auf die dem DGP zugrunde gelegten Koeffizienten über alle 20 bzw. 200 Replikationen des DGP bezeichnet. $\overline{Rmse}$ beinhaltet dementsprechend neben den einzelnen Verzerrungen ($Bias$) bzgl. der wahren Parameter auch die Standardabweichungen ($Stab$) der Schätzwerte über alle Replikationen des DGP. Neben $Bias$, $Stab$ und $Rmse$ werden auch jeweils das arithmetische Mittel ($\bar{\hat{\theta}}$), das Minimum (Min), der Median (Med), sowie das Maximum (Max) der SMLM-Schätzwerte über alle 20 bzw. 200 Replikationen des DGP ausgewiesen. Bei den ersten beiden Experimenten werden darüber hinaus die 25%- bzw. 75%-Quantile (25%, 75%) über alle 200 Replikationen des DGP betrachtet.

5.2.1 Erstes Experiment: Einperiodiges Vieralternativen-Probitmodell

In diesem einperiodigen Probitmodell ist bei den Übersichtsstatistiken in Tabelle 1 zu erkennen, daß erwartungsgemäß bei gleichem N und bei einer Zunahme von $R = 10$ auf $R = 50$ $\overline{|Bias|}$ sowie bei gleichem R und mit zunehmenden N $\overline{Rmse}$ sinkt. Falls dagegen die Anzahl der Simulationsreplikationen von $R = 50$ auf $R = 200$ (bei $N = 1000$) steigt, erhöht sich $\overline{|Bias|}$. Ebenso steigt $\overline{Rmse}$ bei wachsendem R, d.h. mit einer zunehmenden Anzahl an Simulationsreplikationen erhöht sich im Durchschnitt die Streuung der Parameterschätzungen. Bei $R = 50$ steigt zudem leicht $\overline{|Bias|}$ bei einer Erhöhung von $N = 1000$ auf $N = 2000$.

Eine bessere Einsicht in die SMLM erhält man bei der umfassenden Betrachtung der Schätzwerte der einzelnen Parameter. In Tabelle 2 sind ausführliche Statistiken zur SMLM-Schät-

Tabelle 1: Zusammenfassende Statistiken bei der Schätzung im einperiodigen Vieralternativen-Probitmodell

	$\overline{\lvert Bias \rvert}$	$\overline{Rmse}$
$N = 1000$ $R = 10$	0.0430	0.1756
$N = 1000$ $R = 50$	0.0209	0.1797
$N = 1000$ $R = 200$	0.0267	0.1799
$N = 2000$ $R = 10$	0.0320	0.1303
$N = 2000$ $R = 50$	0.0249	0.1368

Tabelle 2: Schätzung der Parameter der erklärenden Variablen im einperiodigen Vieralternativen-Probitmodell

	θ	$\bar{\hat{\theta}}$	$Bias$	Min	25%	Med	75%	Max	$Stab$	$Rmse$
$N = 1000$	γ_1	1.0208	0.0208	0.8176	0.9579	1.0129	1.0782	1.2778	0.0888	0.0912
$R = 10$	γ_2	0.0006	0.0006	-0.0537	-0.0104	-0.0009	0.0117	0.0471	0.0169	0.0169
$N = 1000$	γ_1	1.0167	0.0167	0.8325	0.9533	1.0060	1.0693	1.3007	0.0884	0.0900
$R = 50$	γ_2	0.0007	0.0007	-0.0458	-0.0102	-0.0022	0.0113	0.0487	0.0162	0.0162
$N = 1000$	γ_1	1.0145	0.0145	0.8395	0.9536	1.0087	1.0673	1.2938	0.0872	0.0884
$R = 200$	γ_2	0.0007	0.0007	-0.0501	-0.0106	-0.0018	0.0109	0.0489	0.0162	0.0162
$N = 2000$	γ_1	1.0066	0.0066	0.8917	0.9617	0.9994	1.0343	1.1836	0.0560	0.0564
$R = 10$	γ_2	0.0000	0.0000	-0.0420	-0.0078	0.0001	0.0076	0.0290	0.0112	0.0112
$N = 2000$	γ_1	1.0011	0.0011	0.8915	0.9627	0.9947	1.0373	1.1975	0.0539	0.0539
$R = 50$	γ_2	-0.0002	-0.0002	-0.0397	-0.0078	0.0002	0.0067	0.0269	0.0109	0.0109

zung der Parameter der (alternativenspezifischen) erklärenden Variablen abgebildet. Zu erkennen ist, daß man schon bei (moderatem) $N = 1000$ und (kleinem) $R = 10$ stabile Parameterschätzungen mit sehr geringen Verzerrungen erhält. Nicht nur die arithmetischen Mittel, sondern auch die jeweiligen Mediane über alle Replikationen des DGP liegen sehr nahe an den entsprechenden (wahren) Parametern des DGP. Ausnahmslos ist dabei (erwartungsgemäß) die Güte der SMLM-Schätzung von γ_2 deutlich höher. Wegen des sinkenden $Stab$ vermindert sich mit wachsendem N bei γ_1 und bei γ_2 auch $Rmse$, durchweg auf niedrigem Niveau. Hinsichtlich empirischer Arbeiten deuten diese Ergebnisse an, daß die Parameter der erklärenden Variablen im einperiodigen Mehralternativen-Probitmodell bei korrekter Modellspezifikation schon bei einem moderaten Beobachtungsumfang N und bei einer geringen Anzahl R an Simulationsreplikationen mit der SMLM sehr präzise und stabil geschätzt werden können.

Tabelle 3: Schätzung der Varianz-Kovarianz-Parameter im einperiodigen Vieralternativen-Probitmodell, $N = 2000$, $R = 50$

θ	$\bar{\hat{\theta}}$	$Bias$	Min	25%	Med	75%	Max	$Stab$	$Rmse$
σ_{η_1}	1.4758	-0.0242	1.1173	1.3613	1.4700	1.5735	1.9888	0.1453	0.1473
σ_{η_2}	0.4515	-0.0485	0.0110	0.3104	0.4617	0.6196	1.0207	0.2225	0.2277
$corr(\eta_{i1t}, \eta_{i2t})$	0.4891	-0.0109	-0.6902	0.3251	0.4895	0.6309	0.9984	0.2329	0.2332
$corr(\eta_{i1t}, \eta_{i3t})$	0.4835	-0.0165	0.2114	0.4195	0.4881	0.5481	0.6667	0.0894	0.0909
$corr(\eta_{i2t}, \eta_{i3t})$	0.5726	0.0726	0.1302	0.4370	0.5388	0.6848	0.9996	0.1799	0.1941

Tabelle 4: Zusammenfassende Statistiken bei der Schätzung im fünfperiodigen Dreialternativen-Probitmodell

	$\overline{\vert Bias \vert}$	$\overline{Rmse}$
$N = 250$ $R = 10$	0.0932	0.2566
$N = 250$ $R = 50$	0.0456	0.2418
$N = 250$ $R = 200$	0.0470	0.2508
$N = 500$ $R = 10$	0.1026	0.2057
$N = 500$ $R = 50$	0.0438	0.1987

Im Gegensatz dazu können die Varianz-Kovarianz-Parameter nicht derart korrekt und stabil geschätzt werden. Beispielhaft sind in Tabelle 3 für $N = 2000$ und $R = 50$ Statistiken für die Schätzung dieser Parameter abgebildet. Zu erkennen sind höhere durchschnittliche Verzerrungen, vor allem bei $corr(\eta_{i2t}, \eta_{i3t})$. Bei kleinerem N bzw. R (die ausführlichen Resultate sind auf Anfrage bei den Autoren erhältlich) erscheinen zum Teil etwas stärkere Verzerrungen bzw. Streuungen der Schätzwerte. Dennoch kann festgehalten werden, daß die Varianz-Kovarianz-Parameter in diesem einperiodigen Mehralternativen-Probitmodell (vor allem in Relation zu den Ergebnissen in den mehrperiodigen Probitmodellen, vgl. unten) mit der SMLM zumindest im Durchschnitt recht gut geschätzt werden können und zwar auch bei kleinerem N bzw. R.

5.2.2 Zweites Experiment: Fünfperiodiges Dreialternativen-Probitmodell

In Tabelle 4 sind zunächst die Übersichtsstatistiken bei der Schätzung im fünfperiodigen Dreialternativen-Probitmodell abgebildet. Dabei zeigt sich hinsichtlich $\overline{|Bias|}$ bzw. $\overline{Rmse}$ ein uneinheitliches Bild. Eindeutig zu erkennen ist lediglich, daß $\overline{Rmse}$ (bei gleichem R) mit zunehmenden Beobachtungsumfang N erwartungsgemäß sinkt. Ähnlich wie im ersten Experiment ist festzuhalten, daß die ausschließliche Erhöhung von R keinen durchweg sen-

Tabelle 5: Schätzung der Parameter der erklärenden Variablen im fünfperiodigen Dreialternativen-Probitmodell

	θ	$\bar{\hat{\theta}}$	$Bias$	Min	25%	Med	75%	Max	$Stab$	$Rmse$
$N = 250$	γ_1	1.0037	0.0037	0.7947	0.9387	1.0040	1.0627	1.1857	0.0835	0.0835
$R = 10$	γ_2	-0.0053	-0.0053	-0.0870	-0.0256	-0.0047	0.0158	0.0752	0.0323	0.0327
$N = 250$	γ_1	0.9964	-0.0036	0.8146	0.9292	0.9900	1.0509	1.3041	0.0898	0.0899
$R = 50$	γ_2	-0.0056	-0.0056	-0.0922	-0.0248	-0.0055	0.0175	0.0722	0.0317	0.0321
$N = 250$	γ_1	0.9977	-0.0023	0.8407	0.9252	0.9950	1.0496	1.3490	0.0894	0.0894
$R = 200$	γ_2	-0.0057	-0.0057	-0.0923	-0.0269	-0.0051	0.0158	0.0710	0.0316	0.0321
$N = 500$	γ_1	0.9972	-0.0028	0.8599	0.9468	0.9952	1.0338	1.1811	0.0608	0.0609
$R = 10$	γ_2	-0.0021	-0.0021	-0.0644	-0.0155	-0.0018	0.0120	0.0600	0.0221	0.0222
$N = 500$	γ_1	0.9893	-0.0107	0.8733	0.9385	0.9854	1.0297	1.1842	0.0610	0.0619
$R = 50$	γ_2	-0.0019	-0.0019	-0.0609	-0.0141	-0.0017	0.0139	0.0592	0.0216	0.0217

kenden Effekt auf $\overline{Rmse}$ besitzt. Ebenfalls analog zum ersten Experiment wächst $\overline{|Bias|}$ (hier für $N = 250$) bei einer Steigerung von $R = 50$ auf $R = 200$. Die starke Verminderung von $\overline{|Bias|}$ bei einer Erhöhung von $R = 10$ auf $R = 50$ ist darauf zurückzuführen, daß bei kleinem $R = 10$ die Autokorrelationsparameter ρ_1 bzw. ρ_2 extrem verzerrt geschätzt werden (vgl. auch unten).

Tabelle 5 enthält die ausführlichen Statistiken zur SMLM-Schätzung der Parameter der erklärenden Variablen. Erneut zeigen sich im Vergleich zum ersten Experiment ähnliche Ergebnisse. Schon bei einem kleinen Beobachtungsumfang ($N = 250$) und bei einer geringen Anzahl an Simulationsreplikationen ($R = 10$) gelangt man zu stabilen Parameterschätzungen mit geringen Verzerrungen, sowohl hinsichtlich der arithmetischen Mittel als auch der Mediane über alle Replikationen des DGP. Zudem wird auch hier γ_2 im Vergleich zu γ_1 mit geringeren Streuungen geschätzt. Somit ergibt sich aus dieser Betrachtung, daß die Parameter der erklärenden Variablen auch im fünfperiodigen Dreialternativen-Probitmodell bei korrekter Modellspezifikation schon bei einem kleinen Beobachtungsumfang N und bei einer geringen Anzahl R an Simulationsreplikationen mit der SMLM sehr genau und stabil geschätzt werden können.

Völlig andere Ergebnisse erhält man dagegen bei der Schätzung der Varianz-Kovarianz-Parameter. In Tabelle 6 sind beispielhaft die Resultate für $N = 500$ und $R = 50$ abgebildet. Im Vergleich zu den Parameterschätzungen der erklärenden Variablen erkennt man hier (und vor allem bei den Koeffizienten der zeitinvarianten stochastischen Effekte) teilweise deutlich größere Verzerrungen sowie instabilere Ergebnisse über die 200 Replikationen des DGP. Zu-

Tabelle 6: Schätzung der Varianz-Kovarianz-Parameter im fünfperiodigen Dreialternativen-Probitmodell, $N = 500$, $R = 50$

θ	$\bar{\hat{\theta}}$	$Bias$	Min	25%	Med	75%	Max	$Stab$	$Rmse$
σ_{η_1}	1.4999	-0.0001	0.7560	1.2379	1.4329	1.6841	2.4187	0.3417	0.3417
$corr(\eta_{i1t}, \eta_{i2t})$	0.4678	-0.0322	0.1481	0.3985	0.4835	0.5379	0.7445	0.1176	0.1220
σ_{α_1}	1.3859	-0.1141	0.0394	1.1927	1.5374	1.6918	2.1840	0.4723	0.4860
σ_{α_2}	0.3843	-0.1157	0.0000	0.1745	0.3949	0.5804	0.8356	0.2316	0.2591
ρ_1	0.7420	-0.0580	0.2471	0.6601	0.7530	0.8591	0.9667	0.1422	0.1536
ρ_2	0.4827	-0.0173	0.0138	0.4075	0.5101	0.5845	0.7186	0.1426	0.1437

Tabelle 7: Zusammenfassende Statistiken bei der Schätzung im achtperiodigen Vieralternativen-Probitmodell

	$\overline{\lvert Bias \rvert}$	$\overline{Rmse}$
$N = 250$ $R = 10$	0.2282	0.3366
$N = 250$ $R = 50$	0.1318	0.2988
$N = 250$ $R = 200$	0.1048	0.2781
$N = 500$ $R = 10$	0.2300	0.2842
$N = 500$ $R = 50$	0.1526	0.2425
$N = 500$ $R = 200$	0.1196	0.2259

dem unterscheiden sich das arithmetische Mittel sowie der Median der Schätzwerte zum Teil (z.B. bei σ_{η_1} und σ_{α_1}) recht deutlich. Bei kleinerem N bzw. R (auch diese Statistiken sind auf Anfrage bei den Autoren erhältlich) erscheinen stellenweise noch deutlich stärkere Verzerrungen bzw. Streuungen der Schätzwerte, insbesondere hinsichtlich der Autokorrelationskoeffizienten. Das heißt, für die korrekte und stabile Schätzung der Autokorrelationsparameter ρ_1 bzw. ρ_2 ist eine hohe Anzahl R an Simulationsreplikationen notwendig. Diese Beobachtung deckt sich mit den Untersuchungen von Geweke u.a. (1997).

5.2.3 Drittes Experiment: Achtperiodiges Vieralternativen-Probitmodell

In Tabelle 7 sind die Übersichtsstatistiken bei der Schätzung im achtperiodigen Vieralternativen-Probitmodell abgebildet. Zu erkennen ist, daß mit zunehmenden R sowohl $\overline{\lvert Bias \rvert}$ als auch $\overline{Rmse}$ sinkt. Auch diese Verminderung wird dadurch verursacht, daß mit wachsendem R die Autokorrelationskoeffizienten mit geringeren Verzerrungen geschätzt werden können. Bei wachsendem N nimmt $\overline{Rmse}$ ebenfalls ab. Analog zum zweiten Experiment steigt auch hier $\overline{\lvert Bias \rvert}$ mit zunehmenden N, allerdings in geringem Ausmaß.

Tabelle 8: Schätzung der Parameter der erklärenden Variablen im achtperiodigen Vieralternativen-Probitmodell

	θ	$\bar{\hat{\theta}}$	$Bias$	Min	Med	Max	$Stab$	$Rmse$
$N = 250$	γ_1	1.0243	0.0243	0.9249	1.0076	1.1882	0.0672	0.0717
$R = 10$	γ_2	1.0222	0.0222	0.8010	1.0062	1.3166	0.1197	0.1219
$N = 250$	γ_1	1.0126	0.0126	0.9078	0.9955	1.1951	0.0705	0.0716
$R = 50$	γ_2	1.0188	0.0188	0.8336	1.0298	1.2650	0.1153	0.1169
$N = 250$	γ_1	1.0213	0.0213	0.9438	0.9944	1.2151	0.0657	0.0692
$R = 200$	γ_2	1.0235	0.0235	0.8693	0.9916	1.2929	0.1056	0.1084
$N = 500$	γ_1	0.9973	-0.0027	0.9435	0.9795	1.0977	0.0460	0.0460
$R = 10$	γ_2	0.9899	-0.0101	0.8691	0.9823	1.1543	0.0706	0.0713
$N = 500$	γ_1	0.9943	-0.0057	0.8933	0.9838	1.1284	0.0583	0.0586
$R = 50$	γ_2	0.9920	-0.0080	0.9047	0.9726	1.1664	0.0680	0.0685
$N = 500$	γ_1	0.9920	-0.0080	0.8945	0.9861	1.1177	0.0537	0.0543
$R = 200$	γ_2	0.9862	-0.0138	0.9061	0.9706	1.1379	0.0617	0.0633

In Tabelle 8 sind die ausführlichen Ergebnisse zur SMLM-Schätzung der Parameter der erklärenden Variablen dargestellt. Hierbei kann die Interpretation der ersten beiden Experimente weitgehend beibehalten werden. Schon bei (kleinem) $N = 250$ und (kleinem) $R = 10$ erhält man stabile Parameterschätzungen mit sehr geringen Verzerrungen. Eine Zunahme von R hat bzgl. $Bias$ kaum Auswirkungen. Wegen des sinkenden $Stab$ vermindert sich auch $Rmse$ bei wachsendem N insbesondere beim Parameter γ_2, aber durchweg auf niedrigem Niveau. Eine sukzessive Erhöhung von R bietet vor allem bzgl. des Parameters γ_1 hinsichtlich $Rmse$ ein uneinheitliches Bild. Festzuhalten ist dabei, daß die Schätzungen von Parameter γ_1 im Vergleich zu den Schätzungen von Parameter γ_2 der Dummy-Variablen geringere Streuungen aufweisen. Im Hinblick auf empirische Arbeiten zeigt sich damit auch in diesem MMPM, daß die Parameter der erklärenden Variablen bei korrekter Modellspezifikation schon bei kleinen Beobachtungsumfängen N und geringen Zahlen R an Simulationsreplikationen sehr genau und stabil geschätzt werden.

Hinsichtlich der Schätzung einzelner Varianz-Kovarianz-Parameter erhält man im Vergleich zum ersten, aber auch zum zweiten Experiment hier die instabilsten Ergebnisse. Beispielhaft sind in Tabelle 9 die ausführlichen Resultate für $N = 500$ und $R = 200$ abgebildet. Wiederum erkennt man deutlich größere Verzerrungen sowie instabilere Resultate als bei der Schätzung der Parameter der erklärenden Variablen. Zudem tauchen auch hier teilweise markante Unterschiede zwischen dem arithmetischen Mittel und dem Median der Schätzwerte

Tabelle 9: Schätzung der Varianz-Kovarianz-Parameter im achtperiodigen Vieralternativen-Probitmodell, $N = 500$, $R = 200$

θ	$\bar{\hat{\theta}}$	$Bias$	Min	Med	Max	$Stab$	$Rmse$
σ_{η_1}	1.4337	-0.0663	1.1468	1.3897	2.0383	0.2263	0.2363
σ_{η_2}	0.3769	-0.1231	0.1261	0.3622	0.8323	0.1995	0.2361
$corr(\eta_{i1t}, \eta_{i2t})$	0.2063	-0.2937	-0.0590	0.1702	0.6621	0.1682	0.3451
$corr(\eta_{i1t}, \eta_{i3t})$	0.3940	-0.1060	0.1001	0.3763	0.7015	0.1449	0.1812
$corr(\eta_{i2t}, \eta_{i3t})$	0.4448	-0.0552	0.2060	0.3791	0.7248	0.1569	0.1668
σ_{α_1}	1.4161	-0.0839	0.0315	1.5106	2.1415	0.4305	0.4390
σ_{α_2}	1.5312	0.0312	1.3768	1.5010	1.7313	0.1056	0.1103
σ_{α_3}	0.3986	-0.1014	0.0921	0.4023	0.8667	0.1863	0.2134
ρ_1	0.7864	-0.0136	0.5231	0.8021	0.9122	0.0958	0.0968
ρ_2	0.1961	-0.6039	-0.4679	0.2205	0.5771	0.2141	0.6555
ρ_3	0.4454	-0.0546	0.1509	0.4589	0.7009	0.1268	0.1386

über alle Replikationen des DGP auf. Bei kleinerem N bzw. R erscheinen erneut insbesondere bei den Autokorrelationskoeffizienten zum Teil noch deutlich verstärkte Verzerrungen (auch diese Statistiken sind auf Anfrage bei den Autoren erhältlich). Die (durchschnittlichen) Auswirkungen insbesondere hinsichtlich *Rmse* lassen sich bei der Betrachtung der Übersichtsstatistiken in Tabelle 7 erkennen. Aber selbst bei derartig hohem $N = 500$ sowie $R = 200$ ist die Schätzung des Korrelationskoeffizienten $corr(\eta_{i1t}, \eta_{i2t})$ und insbesondere des Autokorrelationsparameters ρ_2 im Durchschnitt extrem (nach unten) verzerrt. Vor allem bei der Schätzung von ρ_2 tauchen dabei wiederholt negative Schätzwerte auf. Ein solch ungünstiges Verhalten bei hohem N und R ist bei der Schätzung der Varianz-Kovarianz-Parameter im MMPM mit geringerem J und T (vgl. zweites Experiment) nicht zu erkennen.

5.3 Schlußfolgerungen

Die Monte-Carlo-Studien in diesem Abschnitt zeigen zunächst die Robustheit der SMLM-Schätzung in Verbindung mit dem GHK-Simulator im MMPM. Bei den hier untersuchten Modellspezifikationen, Beobachtungsumfängen N sowie Zahlen R an Simulationsreplikationen entstanden keinerlei Probleme bei der Konvergenz der simulierten Loglikelihoodfunktion zu einem Maximum. Diese Beobachtung ist vor allem im Hinblick auf die in der Literatur beschriebenen Schwierigkeiten bei der Parameterschätzung mit der "method of simulated moments" (vgl. z.B. Mühleisen, 1994, Geweke u.a., 1994, Lee, 1995) zu betonen.

Ein wesentliches Ergebnis der Untersuchungen ist die stabile und im Durchschnitt unverzerrte Schätzung der Parameter der erklärenden Variablen und zwar unabhängig von der betrachteten Probitmodellspezifikation. Sowohl in einperiodigen als auch in mehrperiodigen Mehralternativen-Probitmodellen scheint die SMLM in Verbindung mit dem GHK-Simulator bei korrekter Modellspezifikation hinsichtlich der Schätzung dieser Parameter präzise Ergebnisse zu liefern. Zu betonen ist dabei, daß dieser Sachverhalt bei allen betrachteten Probitmodellen schon bei kleinem Beobachtungsumfang N sowie bei kleiner Anzahl R an Simulationsreplikationen zutrifft.

Im Gegensatz dazu können die Varianz-Kovarianz-Parameter in den untersuchten Mehralternativen-Probitmodellen bei weitem nicht derart stabil und korrekt geschätzt werden. Dabei zeigen sich hinsichtlich der einzelnen Probitmodelle jedoch Unterschiede. Im einperiodigen Vieralternativen-Probitmodell liegen die arithmetischen Mittel bzw. Mediane der geschätzten Varianz-Kovarianz-Parameter über alle Replikationen des DGP relativ nahe am zugrunde gelegten Wert. Dies gilt für alle hier betrachteten Beobachtungsumfänge N sowie Zahlen R an Simulationsreplikationen.

Bei der SMLM/GHK-Schätzung der Varianz-Kovarianz-Parameter in mehrperiodigen Mehralternativen-Probitmodellen spielen dagegen N bzw. R eine viel größere Rolle. Bei kleinem N bzw. R sind zum Teil (gegenüber der ohnehin instabilen Schätzung bei großem N bzw. R) noch deutlich stärkere Verzerrungen bzw. Streuungen der Schätzwerte zu erkennen, insbesondere hinsichtlich der Autokorrelationskoeffizienten. Es zeigt sich somit, daß für die korrekte und stabile Schätzung der Autokorrelationsparameter vor allem eine hohe Anzahl R der Simulationsreplikationen notwendig ist. Diese Beobachtung deckt sich mit den Untersuchungen von Geweke u.a. (1997) im zehnperiodigen Dreialternativen-Probitmodell, aber auch mit denen von Keane (1994), Lee (1995, 1997) oder Hyslop (1999) in binären mehrperiodigen Probitmodellen.

Beim Vergleich zwischen der SMLM/GHK-Schätzung der Varianz-Kovarianz-Parameter im fünfperiodigen Dreialternativen- mit derjenigen im achtperiodigen Vieralternativen-Probitmodell zeigt sich, daß bei ersterer mit großem $N = 500$ und $R = 50$ zumindest im Durchschnitt moderate Verzerrungen vorliegen. Im achtperiodigen Vieralternativen-Probitmodell erscheinen dagegen auch bei großem $N = 500$ und großem $R = 200$ zum Teil extrem starke Verzerrungen. Offenbar spielen die Anzahl J der Kategorien sowie die Anzahl T der Perioden und damit die Größe der einzelnen Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ hinsichtlich der Güte der Schätzung der Varianz-Kovarianz-Parameter im MMPM eine wichtige Rolle.

Die Ergebnisse in mehrperiodigen Mehralternativen-Probitmodellen deuten letztlich an, daß (insbesondere bei kleinem N bzw. R) die Identifikation der Parameter der intertemporalen Verknüpfungen (d.h. der Autokorrelationsparameter einerseits sowie der Parameter der stochastischen Effekte andererseits) schwierig ist. Weitergehende eigene Versuche haben gezeigt, daß die Präzision der Schätzung der Autokorrelationsparameter für den Fall, daß im

DGP keine stochastischen Effekte vorliegen, höher ist. Festzuhalten ist dennoch, daß die hier betrachteten Beobachtungsumfänge N sowie Zahlen R an Simulationsreplikationen zur stabilen und korrekten SMLM/GHK-Schätzung aller Varianz-Kovarianz-Parameter vor allem in mehrperiodigen, teilweise aber auch in einperiodigen Mehralternativen-Probitmodellen nicht ausreichend erscheinen.

Keineswegs darf daraus jedoch gefolgert werden, daß damit die Schätzung der Varianz-Kovarianz-Parameter obsolet wird. Die Idee, die Schätzung der Varianz-Kovarianz-Parameter zu vernachlässigen, klingt zwar insofern attraktiv, als damit Simulationsschätzverfahren wie die SMLM und der damit verbundene Aufwand z.B. durch die MLM-Schätzung im Independent Probitmodell (oder gar im multinomialen Logitmodell) vermieden werden können. Weitergehende eigene Versuche haben jedoch gezeigt, daß die fehlerhafte Vernachlässigung kontemporärer bzw. intertemporaler Korrelationselemente im Rahmen des MMPM zu drastischen systematischen Verzerrungen bei der SMLM/GHK-Schätzung der Parameter der erklärenden Variablen führt (vgl. auch die Untersuchungen von Weeks, 1995, im einperiodigen Dreialternativen-Probitmodell). Das heißt, auch wenn man ausschließlich an den Parametern der erklärenden Variablen interessiert wäre, lohnt sich ein Schätzverfahren, daß die stochastische Struktur eines Probitmodells erfassen kann. In dieser Hinsicht stellt die SMLM in Verbindung mit dem GHK-Simulator für zukünftige Anwendungen eine attraktive Schätzmethodik dar.

Literatur

Alvarez, R. M., Nagler, J. (1997). A New Approach for Modeling Stategic Voting in Multiparty Systems. *Unveröffentlichtes Manuskript.* California Institute of Technology.

Asea, P. K., Turnovsky, S. J. (1998). Capital Income Taxation and Risk-Taking in a Small Open Economy. *Journal of Public Economics* **68** 55-90.

Bertaut, C. C., Starr-McCluer, M. (2000). Household Portfolios in the United States. Erscheint in: *Guiso L. u.a. (Hrsg.). Household Portfolios.* Cambridge, Mass.: MIT Press.

Börsch-Supan, A. (1987). *Econometric Analysis of Discrete Choice. With Applications on the Demand for Housing in the U.S. and West-Germany.* Berlin u.a.

Börsch-Supan, A. (1994). Simulationsmethoden für die Analyse qualitativer Daten. *Allgemeines Statistisches Archiv* **78** 20-39.

Börsch-Supan, A., Hajivassiliou, V. A. (1993). Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models. *Journal of Econometrics* **58** 347-368.

Börsch-Supan, A., Pfeiffer, F. (1992). Determinanten der Selbständigkeit in der Bundesrepublik Deutschland. In: *Hujer R. u.a. (Hrsg.). Herausforderungen an den Wohlfahrtsstaat im strukturellen Wandel.* 257-287, Frankfurt/New York.

Börsch-Supan, A., Hajivassiliou, V. A., Kotlikoff, L. J., Morris, J. N. (1992). Health, Children, and Elderly Living Arrangements. A Multiperiod-Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors. In: *Wise D. A. (ed.). Topics in the Economics of Aging.* 79-104, Chicago.

Bolduc, D. (1992). Generalized Autoregressive Errors in the Multinomial Probit Model. *Transportation Research B* **26B** (2) 155-170.

Bolduc, D., Lacroix, G., Muller, C. (1996). The Choice of Medical Providers in Rural Bénin: A Comparison of Discrete Choice Models. *Journal of Health Economics* **15** 477-498.

Borgers, A., Timmermans, H. (1987). Choice Model Specification, Substitution and Spatial Structure Effects: A Simulation Experiment. *Regional Science and Urban Economics* **17** 29-47.

Bunch, D. S. (1991) Estimability in the Multinomial Probit Model. *Transportation Research B* **25B** (1) 1-12.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *Review of*

Economic Studies **XLVII** 225-238.

CHAMBERLAIN, G. (1984). Panel Data. In: *Griliches, Z., Intrilligator, M. D. (eds.). Handbook of Econometrics, Vol II* 1247-1318.

CHINTAGUNTA, P. K. (1992). Estimating a Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments. *Marketing Science* **11** 386-407.

DANSIE, B. R. (1985). Parameter Estimability in the Multinomial Probit Model. *Transportation Research B* **19B** (6) 526-528.

EYMANN, A., RONNING, G. (1997). Microeconometric Models of Tourists' Destination Choice. *Regional Science and Urban Economics* **27** 735-761.

FOTHERINGHAM, A. S. (1983). A New Set of Spatial Interaction Models: The Theory of Competing Destinations. *Environment and Planning* **A15** 15-36.

GEWEKE, J., KEANE, M., RUNKLE, D. (1994). Alternative Computational Approaches to Inference in the Multinomial Probit Model. *The Review of Economics and Statistics* **LXXVI** (4) 609-632.

GEWEKE, J., KEANE, M., RUNKLE, D. (1997). Statistical Inference in the Multinomial Multiperiod Probit Model. *Journal of Econometrics* **80** 125-165.

GÖNÜL, F., SRINIVASAN, K. (1993), Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues. *Marketing Science* **12** 213-229.

GOURIÉROUX, C., MONFORT, A. (1996). *Simulation-Based Econometric Methods.* Oxford University Press.

GUADAGNI, P. M., LITTLE, J. D. C. (1983). A Logit Model of Brand Choice Calibrated on Scanner Data. *Marketing Science* **2** 203-238.

HAJIVASSILIOU, V. A., RUUD, P. (1994). Classical Estimation Methods for LDV Models Using Simulation. In: *Engle, R. F., McFadden, D. (eds.). Handbook of Econometrics, Vol IV* 2383-2441.

HAJIVASSILIOU, V. A., MCFADDEN, D., RUUD, P. (1996). Simulation of Multivariate Normal Rectangle Probabilities and their Derivations. Theoretical and Computational Results. *Journal of Econometrics* **72** 85-134.

HARRIS, K. M., KEANE, M. (1999). A Model of Health Plan Choice: Inferring Preferences and Perceptions from a Combining of Revealed Preference and Attitudinal Data. *Journal of Econometrics* **89** 131-157.

HECKMAN, J. J. (1981). Statistical Models for Discrete Panel Data. In: *Manski, C.*

F., McFadden, D. (eds.). Structural Analysis of Discrete Data with Econometric Applications. 114-178, Cambridge/London.

Hyslop, D. R. (1999). State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. *Econometrica* **67** (6) 1255-1294.

Inkmann, J. (1999). Misspecified Heteroscedasticity in the Panel Probit Model: A Small Sample Comparison of GMM and SML Estimation. *CoFE Discussion Paper No. 99/04.* Universität Konstanz.

Jain, D., Vilcassim, N., Chintagunta, P. (1994). A Random Coefficients Logit Brand Choice Model Applied to Panel Data. *Journal of Business and Economic Statistics* **12** 317-328.

Kaltenborn, U. (1997). Die Anwendung simulativer Schätzverfahren für Discrete-Choice-Modelle am Beispiel von Daten des Sozioökonomischen Panels. *Dissertation.* Fachbereich Wirtschaftswissenschaft der Freien Universität Berlin.

Keane, M. (1994). A Computationally Practical Simulation Estimator for Panel Data. *Econometrica* **62** (1) 95-116.

Lee, L.-F. (1995). Asymptotic Bias in Simulated Maximum Likelihood Estimation of Discrete Choice Models. *Econometric Theory* **11** 437-483.

Lee, L.-F. (1997). Simulated Maximum Likelihood Estimation of Dynamic Discrete Choice Statistical Models. Some Monte Carlo Results. *Journal of Econometrics* **82** 1-35.

Lerman, S. R., Manski, C. F. (1981). On the Use of Simulated Frequencies to Approximate Choice Probabilities. In: *Manski, C. F., McFadden, D. (eds.). Structural Analysis of Discrete Data with Econometric Applications.* 305-319, Cambridge/London.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics.* Cambridge.

Maier, G., Weiss, P. (1990). *Modelle diskreter Entscheidungen. Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften.* Springer-Verlag Wien/New York.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In: *Zarembka P. (ed.). Frontiers in Econometrics.* 105-142, New York.

McFadden, D. (1978). Modelling the Choice of Residential Location. In: *Karlqvist, Lundqvist, A. L., Snickars, F., Weibull, J. W. (eds.). Spatial Interaction Theory and Planning Models.* 75-96, Amsterdam: North-Holland.

McFadden, D. (1984). Econometric Analysis of Qualitative Response Models. In: *Griliches, Z., Intriligator, M. D. (eds.). Handbook of Econometrics, Vol. II* 1395-1457.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica* **57** (5) 995-1026.

McFadden, D., Train, K. (2000). Mixed MNL Models for Discrete Response. Erscheint in: *Journal of Applied Econometrics.*

Mühleisen, M. (1994). *Human Capital Decay and Persistence. A Simulation Approach to German Unemployment.* Campus Verlag Frankfurt am Main.

Ronning, G. (1991). *Mikroökonometrie.* Springer-Verlag Berlin u.a.

Stern, S. (1999). Simulation Based Inference in Econometrics: Motivation and Methods. Erscheint in: *Mariano R. u.a. (eds.). Simulation-Based Inference in Econometrics: Methods and Applications.* Cambridge.

Vijverberg, W. P. M. (1997). Monte Carlo Evaluation of Multivariate Normal Probabilities. *Journal of Econometrics* **76** 281-307.

Weeks, M. (1995). Circumventing the Curse of Dimensionality in Applied Work Using Computer Intensive Methods. *The Economic Journal* **105** 520-530.

Wilde, J. (1999). Gemischte simultane Modelle für Querschnittsdaten. *Dissertation.* Wirtschaftswissenschaftliche Fakultät der Martin-Luther-Universität Halle-Wittenberg.

Ziegler, A. (2000). Simulierte klassische Parameterschätzung in Probitmodellen. *Discussion Paper Nr. 578-00.* Institut für Volkswirtschaftslehre und Statistik, Universität Mannheim.