Institut für
Volkswirtschaftslehre
und Statistik

No. 578-00

Simulierte klassische Parameterschätzung in Probitmodellen

Andreas Ziegler

Beiträge zur angewandten Wirtschaftsforschung

Universität Mannheim
A5, 6
D-68131 Mannheim

Simulierte klassische Parameterschätzung in Probitmodellen

Andreas Ziegler
Seminar für Statistik

11. April 2000

1 Einleitung

In der empirischen Wirtschaftsforschung war die Parameterschätzung in Probitmodellen mit der Maximum-Likelihood-Methode (MLM) und der Verallgemeinerten Momentenmethode (GMM) in der Vergangenheit auf bestimmte Spezifikationen beschränkt. So wurde die MLM vor allem bei einfach strukturierten bzw. einperiodigen Probitmodellen eingesetzt (vgl. z.B. Hausman/Wise, 1978, Ronning, 1991). Die frühere Konzentration auf einperiodige diskrete Entscheidungsmodelle hängt auch damit zusammen, daß kaum entsprechende Paneldaten, bei denen qualitative Variablen über mehrere Perioden beobachtet werden, vorlagen. Mit der zunehmenden Verfügbarkeit derartiger Datensätze können und sollten aber zeitliche Aspekte berücksichtigt werden, da viele ökonomische Entscheidungsprozesse durch komplexe intertemporale Abhängigkeiten beeinflußt sind. Nun wurde mit der Entwicklung der GMM (vgl. z.B. Hansen, 1982, Newey, 1990, 1993) die Schätzung im binären mehrperiodigen Probitmodell ohne strenge intertemporale Restriktionen ermöglicht (vgl. Lechner/Breitung, 1996, Bertschek/Lechner, 1998, Inkmann, 1999). Bei der Untersuchung vieler ökonomischer Fragestellungen ist es jedoch sinnvoll, Mehralternativen-Probitmodelle zu verwenden. Beispiele sind die Analyse der Wohnungsnachfrage älterer Personen (vgl. z.B. Börsch-Supan u.a., 1992), der Produktmarkenwahl von Konsumenten (vgl. z.B. Chintagunta, 1992), der Nachfrage nach verschiedenen medizinischen Behandlungsformen (vgl. z.B. Bolduc u.a., 1996), der Form des Arbeitsstatus verheirateter Frauen (vgl. z.B. Weeks, 1997) sowie der Portfoliowahl von Haushalten (vgl. z.B. Asea/Turnovsky, 1998).

Die Parameterschätzung in einem allgemein formulierten Mehralternativen-Mehrperioden-Probitmodell (MMPM) ist aber ohne restriktive Verteilungsannahmen grundsätzlich weder mit der MLM noch mit der GMM möglich. Dies liegt darin begründet, daß eine derartige Schätzung wegen auftauchender Mehrfachintegrale rechnerisch nicht handhabbar ist. Erst durch die Entwicklung von Simulationsmethoden (vgl. z.B. die Übersichten in Hajivassiliou u.a., 1996 oder Vijverberg, 1997) ist man in der Lage, diese Vielfachintegrale schnell und genau zu approximieren. Mit der Verknüpfung solcher Simulationsmethoden mit klassischen Schätzverfahren wurden in der Literatur verschiedene simulierte klassische Schätzer entwickelt und diskutiert (vgl. z.B. Lerman/Manski, 1981, McFadden, 1989, Börsch-Supan/Hajivassiliou, 1993, Börsch-Supan, 1994, Hajivassiliou, 1993, Hajivassiliou/Ruud, 1994, Keane, 1994, Lee, 1992, 1995, Gourieroux/Monfort, 1991, 1993, 1996, Hajivassiliou/McFadden, 1998).

Im Hinblick auf empirische Untersuchungen werden in dieser Arbeit die Simulierte Maximum-Likelihood-Methode (SMLM), die Simulierte Verallgemeinerte Momentenmethode (SGMM) sowie die Methode der Simulierten Scores (MSS) systematisch im Rahmen des MMPM erläutert. Damit gewinnt man hinsichtlich des Einsatzes des MMPM einen kompakten Überblick über derartige Schätzer, die bisher in der Literatur über viele Artikel verstreut betrachtet werden (vgl. oben). Ausführlich wird die auf der Grundlage der GMM entwickelte SGMM diskutiert. Bei der Ableitung von SGMM- (bzw. GMM-) Schätzern ist insbesondere die konkrete Ausgestaltung der (bedingten) Momentenfunktionen entscheidend. In dieser Arbeit werden zwei denkbare Formulierungen solcher Momentenfunktionen in Probitmodellen gegenüber gestellt. Da sich die Darstellung auf ein allgemein formuliertes MMPM bezieht, werden die bisher in der Literatur lediglich für spezielle Probitmodelle vorgeschlagenen Momentenrestriktionen im Rahmen der GMM bzw. SGMM (vgl. z.B. McFadden, 1989, Geweke u.a., 1994, Bertschek/Lechner, 1998, Inkmann, 1999, Stern, 1999) erweitert. Für die Parameterschätzung im MMPM können somit die vorgeschlagenen Versionen der SGMM direkt implementiert werden.

Das wesentliche Anliegen dieser Arbeit ist die Darstellung der besonderen statistischen Beziehungen zwischen den betrachteten simulierten klassischen Schätzverfahren im MMPM. Eine derartige systematische Diskussion der Interdependenzen zwischen diesen Schätzern speziell in Probitmodellen wurde bislang in der Literatur weitgehend vernachlässigt. Die vergleichende Analyse bezieht sich zunächst direkt auf die einzelnen Schätzansätze. Dabei wird insbesondere (auf der Grundlage des Zusammenhangs zwischen der MLM und der GMM) die spezielle Beziehung zwischen der SMLM und der SGMM erläutert. Es wird aber auch herausgearbeitet, inwiefern die MSS eine ganze Klasse von Schätzverfahren darstellt, die die SMLM sowie bestimmte Versionen der SGMM als Spezialfälle enthält. Die vergleichende Betrachtung in dieser Arbeit richtet sich jedoch vor allem auf die in der Literatur beschriebenen asymptotischen Eigenschaften der betrachteten Schätzer. Hinsichtlich der Asymptotik spielt bei der SGMM (bzw. auch bei der GMM) die Ausgestaltung der Instrumentenmatrix eine zentrale Rolle. Durch die Beleuchtung des Zusammenhangs zwischen der SMLM und der SGMM wird die optimale Formulierung der Instrumentenmatrix speziell im MMPM abgeleitet. Daran anknüpfend wird eine in der Literatur vorgeschlagene (mehrstufige) Schätzstrategie erläutert und kritisch hinterfragt. Letztlich wird argumentiert, warum wir im Hinblick auf empirische Arbeiten nach der Abwägung aller Vor- und Nachteile der betrachteten Schätzverfahren zur Parameterschätzung in einem allgemein formulierten MMPM die SMLM vorschlagen.

Die vorliegende Arbeit ist wie folgt strukturiert. Der zweite Abschnitt beschreibt zunächst einen allgemeinen Ansatz des MMPM und weist dabei auf die Problematik auftauchender Vielfachintegrale hin. Im dritten und vierten Abschnitt wird im Rahmen des MMPM auf der Grundlage der MLM die SMLM sowie auf der Grundlage der GMM die SGMM beschrieben. Im fünften Abschnitt wird dann anknüpfend an die im dritten und vierten Abschnitt diskutierten asymptotischen Eigenschaften der SMLM und der SGMM der Zusammenhang zwischen diesen beiden Schätzverfahren speziell im MMPM beleuchtet. Der sechste Abschnitt erläutert die MSS, die Ideen der SMLM und der SGMM miteinander verknüpft. Im siebten Abschnitt nehmen wir abschließend zu praktischen Aspekten simulierter klassischer Schätzverfahren Stellung.

2 Mehrperioden-Mehralternativen-Probitmodelle

Ausgangspunkt der ökonomischen Ableitung diskreter Mehrperioden-Mehralternativen-Entscheidungsmodelle ist die Möglichkeit einer Untersuchungseinheit i, zum Zeitpunkt t unter J verschiedenen Alternativen einer qualitativen Variablen zu wählen. Beispielsweise läßt sich folgende hypothetische Nutzenfunktion zugrunde legen:

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

Dabei gliedern sich die erklärenden Variablen in einen K_1-dimensionalen Vektor $x_{it} = (x_{it1}, \ldots, x_{itK_1})'$ individuenspezifischer 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.

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

$$D_{ijt} = \begin{cases} 1 & \text{falls Beobachtungseinheit } i \text{ zum Zeitpunkt } t \text{ Kategorie } j \text{ wählt} \\ 0 & \text{sonst} \end{cases} \tag{2}$$

Je nach der Verteilungsannahme bzgl. der ε_{ijt} gelangt man zu unterschiedlichen diskreten Entscheidungsmodellen. Probitmodelle gründen auf folgender Annahme:

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

Daraus folgt, daß auch Differenzen unbeobachtbarer stochastischer Nutzenkomponenten gemeinsam normalverteilt sind mit Erwartungsvektor 0.

Im folgenden werden alle (freien) Parameter des jeweils betrachteten Probitmodells im Vektor θ zusammengefaßt. Gemäß der stochastischen Nutzenmaximierungshypothese (vgl. z.B. Börsch-Supan, 1987, S. 12 ff) entscheidet sich Untersuchungseinheit i zum Zeitpunkt t für Kategorie j, falls $\upsilon_{ijt} > \upsilon_{ikt}$ ($\forall k \neq j; k, j = 1, \ldots, J; i = 1, \ldots, N; t = 1, \ldots, T$). Daraus lassen sich die Wahrscheinlichkeiten $P_{ijt}(\theta)$, daß Untersuchungseinheit i zu einem bestimmten Zeitpunkt t eine Kategorie j wählt, ableiten. Diese Auswahlwahrscheinlichkeiten sind im allgemeinen (ohne vereinfachende Verteilungsannahmen bzgl. der ε_{ijt}) durch $(J-1)$-dimensionale Integrale gekennzeichnet (vgl. Eymann/Ziegler, 2000).

Im Zeitablauf steht jede Untersuchungseinheit i vor der Wahl zwischen J^T verschiedenen potentiellen Kategoriensequenzen. Dementsprechend muß sich eine Beobachtungseinheit i im Hinblick auf die tatsächlich gewählte Kategoriensequenz s in jeder Periode t für eine bestimmte Kategorie j_{it} (mit $j_{it} = 1, \ldots, J$) entscheiden. Die entsprechende Wahrscheinlichkeit $P_{is}(\theta)$ für die Auswahl einer bestimmten Sequenz s durch Untersuchungseinheit i ist im allgemeinen (ohne vereinfachenden Verteilungsannahmen bzgl. der ε_{ijt}) durch $(J-1) \cdot T$-dimensionale Integrale gekennzeichnet (vgl. Eymann/Ziegler, 2000).

Bei großen J und/oder T lassen sich diese durch Vielfachintegrale gekennzeichneten Auswahlwahrscheinlichkeiten mit herkömmlichen (deterministischen) numerischen Integrationsmethoden nicht berechnen. Stattdessen können die Wahrscheinlichkeiten aber schnell und genau mit Hilfe von (stochastischen) Simulationsmethoden, d.h. mit R wiederholten transformierten Ziehungen von Pseudo-Zufallszahlen, approximiert werden. Eine Diskussion verschiedener (erwartungstreuer) Simulationsverfahren findet sich z.B. in Hajivassiliou u.a., 1996, Vijverberg, 1997 oder Wilde, 1999. Damit gelangt man zu den simulierten Auswahlwahrscheinlichkeiten $\widetilde{P}_{ijt}(\theta)$ bzw. $\widetilde{P}_{is}(\theta)$. Diese Simulatoren lassen sich dann mit verschiedenen Schätzverfahren verknüpfen.

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

$$Y_{is} = \begin{cases} 1 & \text{falls Untersuchungseinheit } i \text{ die Kategoriensequenz } s \text{ wählt} \\ 0 & \text{sonst} \end{cases} \tag{3}$$

Dabei gilt $s \in S$, wobei S die Menge aller J^T potentiellen Kategoriensequenzen darstellt. Der wahre unbekannte und zu schätzende Parametervektor wird speziell mit $\dot{\theta}$ bezeichnet. Mit $f(y_i|x_i;\theta)$ wird im folgenden die (bedingte) Masse- oder Dichtefunktion von Y_i gegeben X_i symbolisiert.

3 Simulierte Maximum-Likelihood-Methode

3.1 Ausgangspunkt: Maximum-Likelihood-Methode

Grundlage der Simulierten Maximum-Likelihood-Methode (SMLM) ist der herkömmliche Schätzer aus der Maximum-Likelihood-Methode (MLM):

$$\hat{\theta}_{MLM} = \arg\max_{\theta} \left[lnL(\theta)\right] = \arg\max_{\theta} \left[\sum_{i=1}^{N} lnf(Y_i|X_i;\theta)\right]$$

Im MMPM ist Y_i (gegeben X_i) multinomialverteilt mit den Parametern 1 und allen $P_{is}(\theta)$ $(s \in S)$. Somit ergibt sich im MMPM der MLM-Schätzer:

$$\hat{\theta}_{MLM} = \arg\max_{\theta} \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} lnP_{is}(\theta)\right] \tag{4}$$

Dieser MLM-Schätzer ist eine Lösung folgender Score- bzw. Likelihoodgleichungen:

$$\hat{\theta}_{MLM} = argsolves \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} \frac{\partial lnP_{is}(\theta)}{\partial \theta} = 0\right] \tag{5}$$

Für die Scorefunktionen $S_i(\theta) = \frac{\partial lnf(Y_i|X_i;\theta)}{\partial \theta}$ und speziell im MMPM $S_i(\theta) = \sum_{s \in S} Y_{is} \frac{\partial lnP_{is}(\theta)}{\partial \theta}$ gilt (vgl. Ruud, 1998, S. 376 f):

$$E\left[S_i(\dot{\theta})\right] = 0 \tag{6}$$

Bekanntermaßen liegen für den MLM-Schätzer (unter gewissen Regularitätsbedingungen) hinsichtlich $\dot{\theta}$ Konsistenz und eine asymptotische Normalverteilung vor (vgl. z.B. Amemiya, 1985), d.h. $\sqrt{N}(\hat{\theta}_{MLM} - \dot{\theta}) \xrightarrow{d} NV[0; I(\dot{\theta})^{-1}]$, wobei $I(\dot{\theta})$ die Informationsmatrix bezeichnet. Insbesondere ist die MLM (vor allem im Gegensatz zur GMM, vgl. nächsten Abschnitt) in der Klasse der bzgl. $\dot{\theta}$ konsistenten Schätzer $\hat{\theta}$, bei denen $\sqrt{N}(\hat{\theta} - \dot{\theta})$ gleichförmig über kompakte Teilmengen des Parameterraumes asymptotisch gegen eine Normalverteilung konvergiert, asymptotisch effizient (vgl. z.B. Fomby u.a., 1984, S. 52 ff).

3.2 Einbeziehung von Simulatoren

Falls $f(Y_i|X_i;\theta)$ im MLM-Ansatz durch Mehrfachintegrale gekennzeichnet ist, kann $f(Y_i|X_i;\theta)$ durch einen (unverzerrten) Simulator $\tilde{f}(Y_i,X_i,V_i;\theta)$ substituiert werden. Dabei hängt $\tilde{f}(Y_i,X_i,V_i;\theta)$ im allgemeinen von einem Zufallsterm V_i (d.h. von einem Zufallsvektor oder von einer Zufallsmatrix mit den Zufallsvektoren $V_i^1, V_i^2, \ldots$ als Komponenten) mit bekannter Verteilung ab. Mit den entsprechend dieser Verteilung gezogenen V_{ir} ($i = 1,\ldots,N; r = 1,\ldots,R$) (wobei R die Anzahl der Simulationsreplikationen bezeichnet) gelangt man (vgl. Gouriéroux/Monfort, 1996, S.42) zur Simulierten Maximum-Likelihood-Methode (SMLM) mit dem Schätzer:

$$\hat{\theta}_{SMLM} = \arg\max_{\theta} \left\{ \sum_{i=1}^{N} ln \left[\frac{1}{R} \sum_{r=1}^{R} \tilde{f}(Y_i, X_i, V_{ir}; \theta) \right] \right\} \tag{7}$$

Im MMPM enthält $f(Y_i|X_i;\theta)$ die Auswahlwahrscheinlichkeiten $P_{is}(\theta)$, die bei einer großen Anzahl an Kategorien J und/oder Perioden T aufgrund vorliegender Vielfachintegrale rechnerisch nicht handhabbar sind. Jedoch können die $P_{is}(\theta)$ durch die Simulatoren $\tilde{P}_{is}(\theta)$ approximiert werden (vgl. zweiten Abschnitt). Bei der Verknüpfung des MLM-Ansatzes in (4) mit einem derartigen Simulator $\tilde{P}_{is}(\theta)$ gelangt man zum SMLM-Schätzer im MMPM:

$$\hat{\theta}_{SMLM} = \arg\max_{\theta} \left\{ \sum_{i=1}^{N} \sum_{s \in S} Y_{is} ln \tilde{P}_{is}(\theta) \right\} \tag{8}$$

In Analogie zu (5) ist $\hat{\theta}_{SMLM}$ im MMPM somit eine Lösung folgender simulierter Scoregleichungen:

$$\hat{\theta}_{SMLM} = argsolves \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} \frac{\partial ln \tilde{P}_{is}(\theta)}{\partial \theta} = 0 \right] \tag{9}$$

Im Hinblick auf die numerische Stabilität und die asymptotische Theorie ist für den SMLM-Schätzer zu beachten, daß die Verknüpfung der MLM mit einem Simulator wie z.B. dem Häufigkeitssimulator (vgl. Hajivassiliou u.a., 1996) (bei endlichem R) eine vom Parametervektor θ abhängige diskrete Likelihoodfunktion verursacht. So gelangt man bei derartigen diskreten Simulatoren schon mit einer kleinen Modifikation der Parameterwerte zu einer Veränderung der Zielfunktion in diskreten Schritten. Dies führt einerseits zu numerischen Problemen bei der Optimierung nach θ (vgl. Börsch-Supan/Hajivassiliou, 1993), andererseits zu einer komplexeren asymptotischen Theorie hinsichtlich der simulierten Schätzverfahren. Aus diesem Grund wird bei der Untersuchung der asymptotischen Eigenschaften in diesem (und auch nächsten) Abschnitt davon ausgegangen, daß in die MLM (und in die GMM) ein stetiger Simulator (z.B. der GHK-Simulator, vgl. z.B. Börsch-Supan/Hajivassiliou, 1993, Keane, 1994, Geweke u.a., 1997, Wilde, 1999) eingebettet wird. Zusätzlich sollte beachtet werden, daß die anfänglich erzeugten realisierten Pseudo-Zufallszahlen (Ausgangspunkt sind häufig Realisationen [0;1]-rechteckverteilter Pseudo-Zufallsvariablen) im iterativen Suchprozeß des jeweiligen Schätzers nicht mit den Parametern verändert werden (vgl. Hajivassiliou/Ruud, 1994 bzw. Stern, 1999).

Zu berücksichtigen ist außerdem, daß die $N \cdot R$ gezogenen Zufallsterme V_{ir} über alle Beobachtungseinheiten $i = 1,\ldots,N$ unabhängig sind. Es wird also unterstellt, daß für jede Untersuchungseinheit i unabhängig voneinander jeweils R unterschiedliche Zufallsterme repliziert werden. Damit sind auch die simulierten Auswahlwahrscheinlichkeiten $\widetilde{P}_{is}(\theta)$ im MMPM über alle Beobachtungseinheiten voneinander unabhängig. Im Gegensatz dazu könnten die originär gezogenen Zufallszahlen für alle Untersuchungseinheiten zur Approximation der Auswahlwahrscheinlichkeiten verwendet werden. Im Unterschied zu den im folgenden dargestellten asymptotischen Eigenschaften würde man in diesem Fall zu abweichenden asymptotischen Eigenschaften der SMLM-Schätzer gelangen (vgl. Gouriéroux/Monfort, 1991, Lee, 1992 bzw. 1995).

Im MMPM können zwar die Wahrscheinlichkeiten $P_{is}(\theta)$ durch $\widetilde{P}_{is}(\theta)$ (in Abhängigkeit von den erklärenden Variablen in X_i) erwartungstreu approximiert werden. Allerdings ist die resultierende unverzerrte Simulation der Likelihoodfunktion weder notwendig noch hinreichend für die Konsistenz des SMLM-Schätzers (vgl. Hajivassiliou/Ruud, 1994). Unter einer Reihe von Regularitätsbedingungen läßt sich (z.B. im Rahmen des MMPM) zeigen (vgl. z.B. Lee, 1995, Hajivassiliou/Ruud, 1994, Gouriéroux/Monfort, 1991, 1996), daß der SMLM-Schätzer $\hat{\theta}_{SMLM}$ für eine feste Anzahl R an Simulationsreplikationen und für $N \to \infty$ hinsichtlich $\dot{\theta}$ inkonsistent ist. Diese Inkonsistenz resultiert aus der Wahl von $\widetilde{P}_{is}(\theta)$ als erwartungstreuer Simulator von $P_{is}(\theta)$. Für die Konsistenz wäre es aber notwendig, $lnP_{is}(\theta)$ unverzerrt zu approximieren. Eine derartige erwartungstreue Simulation ist mit stetigen Simulatoren generell unmöglich. Für $R, N \to \infty$ ist $\hat{\theta}_{SMLM}$ dagegen bzgl. $\dot{\theta}$ ein konsistenter Schätzer. Trotz dieser Konsistenz kann bei $\hat{\theta}_{SMLM}$ eine Verzerrung hinsichtlich der asymptotischen Verteilung von $\sqrt{N}(\hat{\theta}_{SMLM} - \dot{\theta})$ vorliegen, falls R im Verhältnis zu N nicht ausreichend schnell gegen unendlich wächst. Lediglich für $\lim_{N\to\infty} \sqrt{N}/R = 0$ ist der SMLM-Schätzer $\hat{\theta}_{SMLM}$ asymptotisch äquivalent mit dem MLM-Schätzer $\hat{\theta}_{MLM}$ und damit asymptotisch effizient (vgl. oben).

4 Simulierte Verallgemeinerte Momentenmethode

4.1 Ausgangspunkt: Verallgemeinerte Momentenmethode

Die Basis der Simulierten Verallgemeinerte Momentenmethode (SGMM) ist der herkömmliche Schätzer aus der Verallgemeinerten Momentenmethode (GMM). Ausgehend von den unabhängigen Beobachtungspaaren (Y_i, X_i) sind (vgl. z.B. Newey, 1993) mit der $(p \times 1)$-dimensionalen Funktion $m(Y_i, X_i; \theta)$ bedingte Momentenrestriktionen ableitbar:

$$E\left[m(Y_i, X_i; \dot{\theta}|X_i)\right] = 0$$

Mit einer Instrumentenmatrix $A(X_i)$, d.h. einer $(l \times p)$-dimensionalen Matrix von Funktionen der erklärenden Variablen in X_i (wobei $l \geq dim\, \theta$) erhält man durch die Minimierung einer quadratischen Form des $(l \times 1)$-dimensionalen Vektors $A(X_i)m(Y_i, X_i; \theta)$ den GMM-Schätzer:

$$\hat{\theta}_{GMM} = \arg\min_{\theta} \left\{ \left[\frac{1}{N}\sum_{i=1}^{N} A(X_i)m(Y_i, X_i; \theta)\right]' W_N \left[\frac{1}{N}\sum_{i=1}^{N} A(X_i)m(Y_i, X_i; \theta)\right] \right\} \tag{10}$$

Dabei stellt W_N eine positiv semidefinite $(l \times l)$-dimensionale Gewichtungsmatrix dar, wobei W_N stochastisch gegen eine beliebige Matrix W konvergiert. Durch die Wahl unterschiedlicher Gewichtungsmatrizen W_N bzw. unterschiedlicher Instrumentenmatrizen $A(X_i)$ gelangt man mit diesem Minimierungsansatz zu verschiedenen GMM-Schätzern $\hat{\theta}_{GMM}$ (vgl. unten). Wenn die Anzahl der (unbedingten) Momentenrestriktionen der Anzahl der zu schätzenden Parameter entspricht (d.h. für $l = dim\theta$) spielt die Wahl von W_N keine Rolle. In diesem Fall erhält man durch das Gleichsetzen von $1/N \sum_{i=1}^{N} A(X_i)m(Y_i, X_i; \theta)$ mit dem Nullvektor den GMM-Schätzer $\hat{\theta}_{GMM}$ in (10).

Bei der Ableitung eines GMM-Schätzers spielt insbesondere die konkrete Ausgestaltung der Funktion $m(Y_i, X_i; \theta)$ eine entscheidende Rolle. Im Rahmen des MMPM ergibt sich mit der Auswahlwahrscheinlichkeit $P_{ijt}(\theta)$ eine denkbare Spezifikation der Momentenrestriktionen. Für die entsprechend (2) bernoulliverteilten Indikatorvariablen D_{ijt}, wobei $P(D_{ijt} = 1| x_{it}, z_{it}; \theta) = P_{ijt}(\theta)$, gilt $E(D_{ijt}|x_{it}, z_{it}; \dot{\theta}) = P_{ijt}(\dot{\theta})$. Daraus folgt $E[(D_{ijt} - P_{ijt}(\dot{\theta}))|x_{it}, z_{it}] = 0$. Mit den $J \cdot T$-dimensionalen Vektoren $D_i = (D_{i11}, \ldots, D_{iJ1}, D_{i12}, \ldots, D_{iJ2}, \ldots\ldots, D_{i1T}, \ldots, D_{iJT})'$ und $P_i(\theta) = [P_{i11}(\theta), \ldots, P_{iJ1}(\theta), P_{i12}(\theta), \ldots, P_{iJ2}(\theta), \ldots\ldots, P_{i1T}(\theta), \ldots, P_{iJT}(\theta)]'$ erhält man (bei strikter Exogenität der erklärenden Variablen, vgl. Lechner/Breitung, 1996) folgende Momentenfunktion:

$$m(Y_i, X_i; \theta) = D_i - P_i(\theta) \tag{11}$$

Damit lautet die erste Variante des GMM-Schätzers im MMPM:

$$\hat{\theta}_{GMM} = \arg\min_{\theta} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(D_i - P_i(\theta)) \right]' W_N \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(D_i - P_i(\theta)) \right] \right\} \tag{12}$$

Bertschek/Lechner (1998) wenden diese Version der GMM zur Parameterschätzung im binären mehrperiodigen Probitmodell an. Da $J = 2$ erfordert die Berechnung der Komponenten $P_{ijt}(\theta)$ des Vektors $P_i(\theta)$ nicht die Lösung mehrdimensionaler Integrale. Im Gegensatz zu dieser Form der GMM steht man bei der MLM in binären mehrperiodigen Probitmodellen vor dem Problem der Berechnung T-dimensionaler Integrale (vgl. oben). Der wesentliche Nachteil dieser Variante der GMM-Schätzung ist die eingeschränkte Möglichkeit der Erweiterung des rechnerisch vereinfachenden Ansatzes auf Mehralternativen-Probitmodelle. Sowohl in mehrperiodigen als auch in einperiodigen Mehralternativen-Probitmodellen müssen bei der Verwendung der Momentenfunktion (11) die Auswahlwahrscheinlichkeiten $P_{ijt}(\theta)$ berechnet werden, die im allgemeinen durch $(J-1)$-dimensionale Integrale gekennzeichnet sind. Damit können bei großem J erneut rechnerische Probleme entstehen, so daß Simulationsverfahren verwendet werden müssen.

Eine alternative Spezifikation von Momentenrestriktionen im MMPM ergibt sich durch die Berücksichtigung der einzelnen Kategoriensequenzen $s \in S$. Für die entsprechend (3) bernoulliverteilten Zufallsvariablen Y_{is}, wobei $P(Y_{is} = 1|X_i; \theta) = P_{is}(\theta)$, gilt $E(Y_{is}|X_i; \dot{\theta}) = P_{is}(\dot{\theta})$. Daraus folgt $E[(Y_{is} - P_{is}(\dot{\theta}))|X_i] = 0$. Mit den J^T-dimensionalen Vektoren $Y_i = (Y_{i1}, Y_{i2}, \ldots)'$ bzw. $P_i^{\bullet}(\theta) = [P_{i1}(\theta), P_{i2}(\theta), \ldots]'$ ergibt sich folgende Momentenfunktion:

$$m(Y_i, X_i; \theta) = Y_i - P_i^{\bullet}(\theta) \tag{13}$$

Damit lautet die zweite Variante des GMM-Schätzers im MMPM:

$$\hat{\theta}_{GMM} = \arg\min_{\theta} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(Y_i - P_i^{\bullet}(\theta)) \right]' W_N \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(Y_i - P_i^{\bullet}(\theta)) \right] \right\} \tag{14}$$

Demnach müssen bei dieser Version der GMM die im allgemeinen durch $(J-1) \cdot T$-dimensionale Integrale gekennzeichneten Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ berechnet werden. Somit ergeben sich schon bei einer moderaten Anzahl an Kategorien J und/oder Perioden T

rechnerische Probleme bei der Parameterschätzung, die selbst im binären Probitmodell bei großem T auftauchen. Bei einperiodigen Probitmodellen gilt $Y_i = D_i$ sowie $P_i(\theta) = P_i^{\bullet}(\theta)$, so daß die Momentenfunktionen in (11) und in (13) ineinander überfließen. In diesem Fall müssen dementsprechend $(J-1)$-dimensionale Integrale berechnet werden.

Ebenso wie beim MLM-Schätzer $\hat{\theta}_{MLM}$ liegen für den GMM-Schätzer $\hat{\theta}_{GMM}$ unter bestimmten Regularitätsbedingungen (vgl. Hansen, 1982) hinsichtlich $\dot{\theta}$ Konsistenz und eine asymptotische Normalverteilung vor. Im Gegensatz zur MLM ist die GMM jedoch nicht in einer allgemeinen Klasse von Schätzverfahren asymptotisch effizient. Zu beachten ist, daß der MLM-Schätzer selbst zur Klasse der GMM-Schätzer gehört. Falls $m(Y_i, X_i; \theta) = S_i(\theta)$ ist entsprechend (6) eine Momentenrestriktion gegeben, so daß der Optimierungsansatz der MLM äquivalent mit dem Minimierungsansatz der GMM ist, falls bei beliebiger Gewichtungsmatrix W_N die Instrumentenmatrix $A(X_i)$ die $(dim\theta \times dim\theta)$-Einheitsmatrix I darstellt (vgl. auch Ruud, 1998).

4.2 Einbeziehung von Simulatoren

Allgemein ausgehend vom herkömmlichen GMM-Schätzer in (10) kann $m(Y_i, X_i; \theta)$ beim Vorliegen von Mehrfachintegralen durch einen (unverzerrten) Simulator $\widetilde{m}(Y_i, X_i, V_i; \theta)$ substituiert werden, wobei $\widetilde{m}(Y_i, X_i, V_i; \theta)$ von einem Zufallsterm V_i mit bekannter Verteilung abhängt. Mit den entsprechend dieser Verteilung wiederholt gezogenen V_{ir} $(i = 1, \ldots, N; r = 1, \ldots, R)$ gelangt man zur Simulierten Verallgemeinerten Momentenmethode (SGMM) mit dem Schätzer (vgl. Gouriéroux/Monfort, 1996, S. 27):

$$\begin{aligned} \hat{\theta}_{SGMM} \quad = \quad \arg\min_{\theta} \Bigg\{ & \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i) \frac{1}{R} \sum_{r=1}^{R} \widetilde{m}(Y_i, X_i, V_{ir}; \theta) \right]' \cdot W_N \cdot \\ & \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i) \frac{1}{R} \sum_{r=1}^{R} \widetilde{m}(Y_i, X_i, V_{ir}; \theta) \right] \Bigg\} \end{aligned} \tag{15}$$

Im MMPM enthält die Momentenfunktion $m(Y_i, X_i; \theta)$ in (11) die Komponenten $P_{ijt}(\theta)$ von $P_i(\theta)$, die analog den obigen Überlegungen mit Hilfe von Simulatoren $\widetilde{P}_{ijt}(\theta)$ approximiert werden. Damit erhält man den Simulator $\widetilde{P}_i(\theta) = [\widetilde{P}_{i11}(\theta), \ldots, \widetilde{P}_{iJ1}(\theta), \ldots\ldots, \widetilde{P}_{i1T}(\theta), \ldots, \widetilde{P}_{iJT}(\theta)]'$ des $J \cdot T$-dimensionalen Vektors $P_i(\theta)$. Bei der Verknüpfung des GMM-Ansatzes in (12) mit einem derartigen Simulator ergibt sich:

$$\begin{aligned} & \hat{\theta}_{SGMM} = \\ & \arg\min_{\theta} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(D_i - \widetilde{P}_i(\theta)) \right]' W_N \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(D_i - \widetilde{P}_i(\theta)) \right] \right\} \end{aligned} \tag{16}$$

In der Literatur wurde diese Variante der SGMM im allgemein formulierten MMPM (nach meiner Kenntnis) noch nicht verwendet bzw. mit der SMLM verglichen. Festzuhalten bleibt, daß bei diesem Schätzansatz die Auswahlwahrscheinlichkeiten $P_{ijt}(\theta)$ für jede Untersuchungseinheit $i = 1, \ldots, N$ und für alle Kategorien $j = 1, \ldots, J$ in sämtlichen Perioden $t = 1, \ldots, T$ zu simulieren sind.

Die zweite Spezifikation der Momentenfunktion $m(Y_i, X_i; \theta)$ im MMPM entsprechend (13) enthält die Komponenten $P_{is}(\theta)$ von $P_i^{\bullet}(\theta)$, die mit Hilfe von Simulatoren $\widetilde{P}_{is}(\theta)$ approximiert werden können. Damit gelangt man zum Simulator $\widetilde{P}_i^{\bullet}(\theta) = [\widetilde{P}_{i1}(\theta), \widetilde{P}_{i2}(\theta), \ldots]'$ des

J^T-dimensionalen Vektors $P_i^\bullet(\theta)$. Der GMM-Ansatz in (14) kann nun mit einem solchen Simulator verknüpft werden, und es folgt:

$$\hat{\theta}_{SGMM} = \tag{17}$$
$$\arg\min_{\theta} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(Y_i - \widetilde{P}_i^\bullet(\theta)) \right]' W_N \left[\frac{1}{N} \sum_{i=1}^{N} A(X_i)(Y_i - \widetilde{P}_i^\bullet(\theta)) \right] \right\}$$

Im Gegensatz zur SMLM, bei der für eine Beobachtungseinheit i lediglich die Wahrscheinlichkeit $P_{is}(\theta)$ für die tatsächlich gewählte Kategoriensequenz s zu approximieren ist, müssen bei dieser Version der SGMM die Wahrscheinlichkeiten $P_{is}(\theta)$ für die Wahl aller J^T potentiellen Kategoriensequenzen $s \in S$ simuliert werden. Bei großen J und/oder T führt dies schnell zu einem immensen Rechenaufwand, so daß diese Variante der SGMM dann nicht handhabbar ist. Einen möglichen Ausweg aus diesen rechnerischen Problemen bietet eine modifizierte SGMM nach Keane (1993, 1994). Durch die Formulierung der Loglikelihoodfunktion im MMPM mit einer Summe logarithmierter (nicht nur bzgl. X_i) bedingter Wahrscheinlichkeiten müssen hier für jede Beobachtungseinheit i anstatt J^T lediglich $J \cdot T$ verschiedene Wahrscheinlichkeiten approximiert werden (vgl. auch Geweke u.a., 1997). Zu beachten ist allerdings, daß sich die asymptotischen Eigenschaften dieses modifizierten SGMM-Schätzers (insbesondere auch hinsichtlich der Konsistenz) von den (im folgenden dargelegten) asymptotischen Eigenschaften des herkömmlichen SGMM-Schätzers unterscheiden.

Der in theoretischer Hinsicht wesentliche Unterschied zwischen der (herkömmlichen) SGMM und der SMLM besteht darin, daß bei der SGMM in den zugrunde liegenden Momentenbedingungen die originären Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ (bzw. $P_{ijt}(\theta)$) linear erscheinen (vgl. McFadden, 1989). Durch die angenommene erwartungstreue Simulation $\widetilde{P}_{is}(\theta)$ von $P_{is}(\theta)$ verschwinden die Simulationsverzerrungen über das Gesetz der großen Zahlen für $N \to \infty$ (vgl. Stern, 1999). Damit ist es im Gegensatz zum SMLM-Schätzer $\hat{\theta}_{SMLM}$ hinsichtlich der Konsistenz des SGMM-Schätzers $\hat{\theta}_{SGMM}$ nicht notwendig, die Anzahl R an Simulationsreplikationen gegen unendlich wachsen zu lassen. Unter einer Reihe von Regularitätsbedingungen ist der SGMM-Schätzer $\hat{\theta}_{SGMM}$ (vgl. McFadden, 1989, Hajivassiliou/Ruud, 1994, Gouriéroux/Monfort, 1991, 1993 und 1996) für eine feste Anzahl R an Simulationsreplikationen und für $N \to \infty$ hinsichtlich $\dot{\theta}$ konsistent. Darüber hinaus liegt für eine feste Anzahl R an Simulationsreplikationen und für $N \to \infty$ eine asymptotische Normalverteilung vor, d.h. $\sqrt{N}(\hat{\theta}_{SGMM} - \dot{\theta}) \xrightarrow{d} NV(0; V_{SGMM})$. V_{SGMM} setzt sich aus zwei additiven Teilen zusammen. Der erste Summand ist die asymptotische Varianz-Kovarianzmatrix V_{GMM}, die sich aus der GMM-Schätzung ergibt. Der zweite Summand beinhaltet den Einfluß der Simulationen. Dabei spielt insbesondere die Varianz des verwendeten Simulators eine wichtige Rolle. Für $R \to \infty$ und $N \to \infty$ besteht zwischen dem SGMM-Schätzer $\hat{\theta}_{SGMM}$ und dem GMM-Schätzer $\hat{\theta}_{GMM}$ Äquivalenz. Diese asymptotischen Eigenschaften bzgl. $\hat{\theta}_{SGMM}$ beziehen sich auf eine jeweilige Klasse von SGMM-Schätzern, für die bestimmte Momentenbedingungen vorliegen. Nun können hier (genauso wie bei der GMM) sowohl die Gewichtungsmatrix W_N als auch die Instrumentenmatrix $A(X_i)$ in verschiedener Weise gewählt werden. Einzelheiten über die optimale Wahl dieser Matrizen bei der SGMM und über den Zusammenhang zur optimalen Wahl dieser Matrizen bei der GMM werden in Gouriéroux/Monfort (1996, S. 31 ff) beschrieben. Letztlich gelangt man für $R \to \infty$ und $N \to \infty$ bei der SGMM aber lediglich zur asymptotischen Varianz-Kovarianzmatrix V_{GMM} von $\sqrt{N}(\hat{\theta}_{GMM} - \dot{\theta})$ der GMM-Schätzung. Die GMM ist jedoch im allgemeinen nicht asymptotisch effizient. Beim SMLM-Schätzer erhält man dagegen mit $\lim_{N\to\infty}\sqrt{N}/R = 0$ den asymptotisch effizienten MLM-Schätzer $\hat{\theta}_{MLM}$. Es stellt sich nun die Frage, inwieweit speziell im MMPM mit der SGMM die asymptotische Effizienz des MLM-Schätzers erreicht werden kann. Zu beachten ist, daß der (oben beschriebene) allgemeine Zusammenhang zwischen der MLM und der GMM nicht direkt auf die Beziehung zwischen der SMLM und der SGMM übertragen werden kann.

5 Zusammenhang zwischen der Simulierten Maximum-Likelihood-Methode und der Simulierten Verallgemeinerten Momentenmethode

5.1 Ausgangspunkt: Zusammenhang zwischen der Maximum-Likelihood-Methode und der Verallgemeinerten Momentenmethode

Im MMPM gilt entsprechend (4) bzw. (5) $\partial lnL(\theta)/\partial\theta = \sum_{i=1}^{N}\sum_{s\in S} Y_{is}[\partial lnP_{is}(\theta)/\partial\theta]$. Darüber hinaus ist im MMPM $\sum_{s\in S} P_{is}(\theta) = 1$ und somit $\partial \sum_{s\in S} P_{is}(\theta)/\partial\theta = \sum_{s\in S}\partial P_{is}(\theta)/\partial\theta = 0$. Zudem gilt $\partial P_{is}(\theta)/\partial\theta = P_{is}(\theta)[\partial lnP_{is}(\theta)/\partial\theta]$. Daraus ergibt sich:

$$\sum_{s\in S} P_{is}(\theta)\frac{\partial lnP_{is}(\theta)}{\partial\theta} = 0$$

Somit erhält man:

$$\begin{aligned}
\frac{\partial lnL(\theta)}{\partial\theta} - \sum_{s\in S} P_{is}(\theta)\frac{\partial lnP_{is}(\theta)}{\partial\theta} &= \sum_{i=1}^{N}\sum_{s\in S} Y_{is}\frac{\partial lnP_{is}(\theta)}{\partial\theta} - \sum_{s\in S} P_{is}(\theta)\frac{\partial lnP_{is}(\theta)}{\partial\theta} \\
\frac{\partial lnL(\theta)}{\partial\theta} &= \sum_{i=1}^{N}\sum_{s\in S}[Y_{is} - P_{is}(\theta)]\,\frac{\partial lnP_{is}(\theta)}{\partial\theta} \\
\frac{\partial lnL(\theta)}{\partial\theta} &= \sum_{i=1}^{N}\frac{\partial lnP_i^{\bullet}(\theta)'}{\partial\theta}\,[Y_i - P_i^{\bullet}(\theta)] \qquad (18)
\end{aligned}$$

Dabei gilt erneut für die J^T-dimensionalen Vektoren $P_i^{\bullet}(\theta) = [P_{i1}(\theta), P_{i2}(\theta), \ldots]'$ und $Y_i = (Y_{i1}, Y_{i2}, \ldots)'$. Durch das Gleichsetzen von (18) mit dem Nullvektor gelangt man zum MLM-Schätzer $\hat{\theta}_{MLM}$. Allerdings leidet dieser Ansatz darunter, daß für jede Untersuchungseinheit $i = 1, \ldots, N$ die Auswahlwahrscheinlichkeiten $P_{is}(\theta)$ für alle Kategoriensequenzen $s \in S$ berechnet werden müssen.

Aufgrund von $E(Y_{is}|X_i;\dot{\theta}) = P_{is}(\dot{\theta})$ bzw. $E(Y_i|X_i;\dot{\theta}) = P_i^{\bullet}(\dot{\theta})$ und $E[Y_i - P_i^{\bullet}(\dot{\theta})|X_i] = 0$ ergeben sich durch $m(Y_i, X_i;\theta) = Y_i - P_i^{\bullet}(\theta)$ entsprechend (13) bedingte Momentenrestriktionen mit $E[m(Y_i, X_i;\dot{\theta}|X_i)] = 0$. Bei der Berücksichtigung einer Instrumentenmatrix $A(X_i)$ lassen sich unbedingte Momentenrestriktionen mit $E[A(X_i)m(Y_i, X_i;\dot{\theta})] = 0$ und somit GMM-Schätzer konstruieren. Die Instrumentenmatrix $A(X_i)$ kann von θ bzw. konkreter von einem (fixierten) Parametervektor $\overline{\theta}$ abhängen, der im iterativen Optimierungsprozeß nicht modifiziert werden darf (vgl. McFadden, 1989). Aus obigem MLM-Ansatz ergibt sich die optimale Instrumentenmatrix $A(X_i)^* = \partial lnP_i^{\bullet}(\dot{\theta})'/\partial\theta$. Dabei wird hinsichtlich der Durchführbarkeit die Instrumentenmatrix mit einem bzgl. $\dot{\theta}$ konsistenten Schätzer $\overline{\theta}$ berechnet, ohne die asymptotischen Eigenschaften des resultierenden GMM-Schätzers zu verändern. Durch diese Spezifikation der Instrumentenmatrix kann man für den GMM-Schätzer Bedingungen erhalten, die den Ableitungen der Loglikelihoodfunktion asymptotisch entsprechen. Damit gelangt man zu einem GMM-Schätzer, der mit dem MLM-Schätzer asymptotisch äquivalent ist. Daraus folgt, daß der GMM-Schätzer bei dieser Spezifikation von $A(X_i)$ auch asymptotisch effizient ist. Die Anzahl der unbekannten Parameter entspricht dann der Anzahl der unbedingten Momentenrestriktionen. Somit erhält man den optimalen GMM-Schätzer im MMPM:

$$\hat{\theta}^*_{GMM} = argsolves \left\{ \sum_{i=1}^{N} \frac{\partial ln P_i^\bullet(\overline{\theta})'}{\partial \theta} \left[Y_i - P_i^\bullet(\theta) \right] = 0 \right\} \tag{19}$$

Zu beachten ist bei dieser (bzgl. der asymptotischen Effizienz interessanten) Schätzstrategie, daß zunächst ein konsistenter Schätzer $\overline{\theta}$ berechnet werden muß. Damit liegt ein zweistufiges Schätzverfahren vor. Hinsichtlich der Praktikabilität in empirischen Anwendungen verliert somit dieser Schätzer insbesondere gegenüber der (einstufigen) MLM an Attraktivität. Bei der Verwendung alternativer Instrumentenmatrizen $A(X_i)$ oder anderer Funktionen $m(Y_i, X_i; \theta)$ liegen für den GMM-Schätzer zwar hinsichtlich $\dot{\theta}$ Konsistenz und eine asymptotische Normalverteilung vor, aber keine asymptotische Effizienz in einer allgemeinen Klasse von Schätzverfahren (vgl. Bertschek/Lechner, 1998, für das binäre mehrperiodige Probitmodell).

5.2 Einbeziehung von Simulatoren

Übertragen auf die SMLM und die SGMM im MMPM ergeben sich für den Fall, daß $\sum_{s \in S} \widetilde{P}_{is}(\theta) = 1$ (vgl. Lee, 1995, S. 441) entsprechend (18) und (9) folgende simulierten Scoregleichungen:

$$\begin{aligned} \frac{\partial ln \widetilde{L}(\theta)}{\partial \theta} &= \sum_{i=1}^{N} \sum_{s \in S} \frac{\partial ln \widetilde{P}_{is}(\theta)}{\partial \theta} \left[Y_{is} - \widetilde{P}_{is}(\theta) \right] \\ &= \sum_{i=1}^{N} \frac{\partial ln \widetilde{P}_i^\bullet(\theta)'}{\partial \theta} \left[Y_i - \widetilde{P}_i^\bullet(\theta) \right] \end{aligned} \tag{20}$$

Dabei gilt wieder für die J^T–dimensionalen Vektoren $\widetilde{P}_i^\bullet(\theta) = [\widetilde{P}_{i1}(\theta), \widetilde{P}_{i2}(\theta), \ldots]'$ und $Y_i = (Y_{i1}, Y_{i2}, \ldots)'$. Ausgehend von dieser Überlegung gelangt man in Anlehnung an (17) mit einer $(dim\, \theta \times J^T)$-dimensionalen Instrumentenmatrix $A(X_i)$ durch das Gleichsetzen von

$$\sum_{i=1}^{N} A(X_i) \left[Y_i - \widetilde{P}_i^\bullet(\theta) \right]$$

mit dem Nullvektor zu einem SGMM-Schätzer. Allerdings kann die Matrix $\partial ln \widetilde{P}_i^\bullet(\theta)'/\partial \theta$ aus der SMLM, berechnet mit einem (hinsichtlich $\dot{\theta}$) konsistenten Schätzer $\overline{\theta}$, nicht direkt übernommen werden. Zu beachten ist, daß selbst $\partial ln \widetilde{P}_i^\bullet(\dot{\theta})'/\partial \theta$ mit den Komponenten $\partial ln \widetilde{P}_{is}(\dot{\theta})/\partial \theta = \partial ln \widetilde{P}_{is}(V_{i1}^s, \ldots, V_{iR}^s; \dot{\theta})/\partial \theta$, wobei $V_{ir}^s = (V_{i1r}^s, \ldots, V_{iHr}^s)'$ $(r = 1, \ldots, R)$ gilt, keine Instrumentenmatrix darstellt (vgl. Gouriéroux/Monfort, 1996, S. 95). Durch die Zufallsvektoren V_{ir}^s, die sowohl in $\partial ln \widetilde{P}_{is}(\dot{\theta})/\partial \theta$ als auch in $\widetilde{P}_{is}(\dot{\theta})$ im Rahmen von $[Y_{is} - \widetilde{P}_{is}(\dot{\theta})]$ eingehen, entsteht eine Korrelation zwischen $\partial ln \widetilde{P}_{is}(\dot{\theta})/\partial \theta$ und $[Y_{is} - \widetilde{P}_{is}(\dot{\theta})]$. Damit liegt eine

Korrelation zwischen $\partial ln\widetilde{P}_{is}(\theta)/\partial\theta$ und $[Y_{is} - \widetilde{P}_{is}(\theta)]$ vor, selbst wenn diese Größen mit Hilfe eines konsistenten Schätzers $\overline{\theta}$ simuliert werden. Mit anderen Worten entspricht

$$E\left\{\frac{\partial ln\widetilde{P}_{is}(\dot{\theta})}{\partial\theta}\left[Y_{is} - \widetilde{P}_{is}(\dot{\theta})\right]\right\}$$

nicht notwendigerweise dem Nullvektor. Um diese Korrelation zu beseitigen, können die $\partial ln\widetilde{P}_{is}(\theta)/\partial\theta$ mit alternativen Zufallsvektoren $\ddot{V}_{ir}^s = (\ddot{V}_{i1r}^s, \ldots, \ddot{V}_{iHr}^s)'$ $(r = 1, \ldots, \ddot{R})$ simuliert werden. Die Voraussetzung für das Vorliegen einer Instrumentenmatrix ist dabei die Unabhängigkeit der $\ddot{V}_{ir}^s$ und V_{ir}^s. Damit ist $\partial ln\widetilde{P}_{is}(\ddot{V}_{i1}^s, \ldots, \ddot{V}_{i\ddot{R}}^s; \dot{\theta})/\partial\theta$ unabhängig von $\widetilde{P}_{is}(V_{i1}^s, \ldots, V_{iR}^s; \dot{\theta})$, so daß gilt:

$$E\left\{\frac{\partial ln\widetilde{P}_{is}(\ddot{V}_{i1}^s, \ldots, \ddot{V}_{i\ddot{R}}^s; \dot{\theta})}{\partial\theta}\left[Y_{is} - \widetilde{P}_{is}(V_{i1}^s, \ldots, V_{iR}^s; \dot{\theta})\right]\right\} = 0$$

Mit Hilfe eines konsistenten Schätzers $\overline{\theta}$ gelangt man im MMPM zum simulierten Analogon von (19), d.h. zu einem SGMM-Schätzer, der die simulierte optimale Instrumentenmatrix enthält:

$$\begin{aligned}\hat{\theta}^*_{SGMM} &= argsolves\left\{\sum_{i=1}^{N}\sum_{s\in S}\frac{\partial ln\widetilde{P}_{is}(\ddot{V}_{i1}^s, \ldots, \ddot{V}_{i\ddot{R}}^s; \overline{\theta})}{\partial\theta}\left[Y_{is} - \widetilde{P}_{is}(V_{i1}^s, \ldots, V_{iR}^s; \theta)\right] = 0\right\} \\ &= argsolves\left\{\sum_{i=1}^{N}\frac{\partial ln\widetilde{P}_i^\bullet(\ddot{V}_{i1}, \ldots, \ddot{V}_{i\ddot{R}}; \overline{\theta})'}{\partial\theta}\left[Y_i - \widetilde{P}_i^\bullet(V_{i1}, \ldots, V_{iR}; \theta)\right] = 0\right\} \quad (21)\end{aligned}$$

Dabei beinhalten $\ddot{V}_{ir}$ $(r = 1, \ldots, \ddot{R})$ bzw. V_{ir} $(r = 1, \ldots, R)$ jeweils die Vektoren $\ddot{V}_{ir}^s$ $(s \in S)$ bzw. V_{ir}^s $(s \in S)$. Zu beachten ist, daß die Komponenten $\partial ln\widetilde{P}_{is}(\theta)/\partial\theta$ von $\partial ln\widetilde{P}_i^\bullet(\theta)'/\partial\theta$ für eine feste Anzahl $\ddot{R}$ an Simulationsreplikationen nicht erwartungstreu approximiert werden können. Wegen der vorliegenden Simulationsverzerrungen muß mit $N \to \infty$ sowohl R als auch $\ddot{R}$ groß sein, um der asymptotischen Effizienz des MLM-Schätzers $\hat{\theta}_{MLM}$ hinreichend nahezukommen (vgl. Gouriéroux/Monfort, 1996, S. 95 bzw. Keane, 1993, S. 549).

Hinsichtlich der asymptotischen Effizienz bietet sich damit für die SGMM-Schätzung im MMPM als Alternative zur SMLM-Schätzung im MMPM bei empirischen Anwendungen folgende mehrstufige Vorgehensweise an (vgl. McFadden, 1989, Stern, 1999). Zunächst ist ein SGMM-Schätzer $\hat{\theta}_{SGMM}$ mit Hilfe einer leicht handhabbaren Instrumentenmatrix abzuleiten. Dieser ist hinsichtlich $\dot{\theta}$ bei einer beliebigen festen Anzahl an Simulationsreplikationen konsistent. Danach kann unter der Verwendung von $\hat{\theta}_{SGMM}$ aus der ersten Stufe die obige optimale Instrumentenmatrix mit einer sehr großen Anzahl $\ddot{R}$ an Simulationsreplikationen approximiert werden. Dies ist rechnerisch weniger aufwendig, da dieser Simulationsvorgang nur einmal durchgeführt werden muß. Schließlich kann mit der approximierten idealen Instrumentenmatrix der optimale SGMM-Schätzer $\hat{\theta}^*_{SGMM}$ berechnet werden.

6 Methode der Simulierten Scores

Die von Hajivassiliou/McFadden (1998) vorgeschlagene Methode der Simulierten Scores (MSS) verbindet Ideen der SMLM und der SGMM. Ausgangspunkt sind die Scorefunktionen aus der MLM mit $\sum_{i=1}^N S_i(\theta) = \sum_{i=1}^N \partial lnf(Y_i|X_i;\theta)/\partial\theta$, wobei man (vgl. auch dritten

Abschnitt) durch die Gleichsetzung des Ausdrucks mit dem Nullvektor zum MLM-Schätzer $\hat{\theta}_{MLM}$ gelangt. Für den Fall, daß die Scorefunktionen wegen vorliegender Vielfachintegrale rechnerisch nicht handhabbar sind, versucht die MSS die $S_i(\theta)$ direkt durch Simulatoren $\widetilde{S}_i(\theta)$ zu approximieren. Damit erhält man allgemein den MSS-Schätzer:

$$\hat{\theta}_{MSS} = argsolves \left[\sum_{i=1}^{N} \widetilde{S}_i(\theta) = 0 \right]$$

Im MMPM konkretisiert sich $\widetilde{S}_i(\theta)$ zu $\sum_{s \in S} Y_{is} \left(\widetilde{\frac{\partial ln P_{is}(\theta)}{\partial \theta}} \right)$, wobei $\widetilde{\frac{\partial ln P_{is}(\theta)}{\partial \theta}}$ einen Simulator für $\partial ln P_{is}(\theta)/\partial\theta$ darstellt. Es ergibt sich:

$$\hat{\theta}_{MSS} = argsolves \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} \left(\widetilde{\frac{\partial ln P_{is}(\theta)}{\partial \theta}} \right) = 0 \right]$$

Diese Definition der MSS beinhaltet eine ganze Klasse denkbarer Schätzer. Beispielsweise können die $S_i(\theta) = \partial ln f(Y_i|X_i;\theta)/\partial\theta$ mit Hilfe einer (erwartungstreuen) Simulation $\widetilde{f}(Y_i, X_i, V_i; \theta)$ für $f(Y_i|X_i;\theta)$ approximiert werden, wobei V_i entsprechend obiger Überlegungen einen bestimmten Zufallsterm mit bekannter Verteilung bezeichnet. Mit den aus dieser Verteilung gezogenen V_{ir} $(i = 1, \ldots, N; r = 1, \ldots, R)$ gelangt man zu folgendem speziellen MSS-Schätzer:

$$\hat{\theta}_{MSS} = argsolves \left\{ \sum_{i=1}^{N} \frac{\partial ln \left[\frac{1}{R} \sum\limits_{r=1}^{R} \widetilde{f}(Y_i, X_i, V_{ir}; \theta) \right]}{\partial \theta} = 0 \right\} \tag{22}$$

Zu erkennen ist, daß dieser MSS-Schätzer entsprechend obiger Überlegungen identisch mit dem SMLM-Schätzer in (7) ist, solange dieselben Simulatoren $\frac{1}{R} \sum_{r=1}^{R} \widetilde{f}(Y_i, X_i, V_{ir}; \theta)$ verwendet werden. Im MMPM ergibt sich entsprechend (8) und (9):

$$\hat{\theta}_{MSS} = argsolves \left[\sum_{i=1}^{N} \sum_{s \in S} Y_{is} \frac{\partial ln \widetilde{P}_{is}(\theta)}{\partial \theta} = 0 \right] \tag{23}$$

Der Zusammenhang zwischen der MSS und der SGMM ergibt sich aus der Überlegung, daß im allgemeinen entsprechend (6) für die Scores $E[S_i(\dot{\theta})] = 0$ gilt. Falls $\widetilde{S}_i(\theta)$ einen unverzerrten Simulator von $S_i(\theta)$ darstellt, ist der resultierende MSS-Schätzer auch ein SGMM-Schätzer (vgl. Gouriéroux/Monfort, 1996, S. 35). Hierin liegt der wesentliche Reiz der MSS. Eine erwartungstreue Simulation von $S_i(\theta)$ würde auch bei einer festen Anzahl R an Simulationsreplikationen zur Konsistenz von $\hat{\theta}_{MSS}$ führen. Darüber hinaus werden durch die direkte Simulation der Scores die im Hinblick auf die asymptotische Effizienz optimalen Instrumente verwendet (vgl. fünften Abschnitt). Somit ist im Rahmen des MMPM der in (21) dargestellte SGMM-Schätzer als MSS-Schätzer zu interpretieren. In diesem Fall werden die $S_i(\theta)$ mit den einzelnen unabhängigen Zufallsvektoren in $\ddot{V}_{ir}$ $(r = 1, \ldots, \ddot{R})$ bzw. in V_{ir} $(r = 1, \ldots, R)$ simuliert:

$$\hat{\theta}_{MSS} = argsolves \left\{ \sum_{i=1}^{N} \frac{\partial ln \widetilde{P}_i^{\bullet}(\ddot{V}_{i1}, \ldots, \ddot{V}_{i\ddot{R}}; \overline{\theta})'}{\partial \theta} \left[Y_i - \widetilde{P}_i^{\bullet}(V_{i1}, \ldots, V_{iR}; \theta) \right] = 0 \right\} \tag{24}$$

Allerdings ist die erwartungstreue Simulation der optimalen Instrumentenmatrix und damit der Scorefunktionen nicht einfach. Hajivassiliou/McFadden (1998) schlagen hierfür den Acceptance-Rejection-Simulator sowie den Gibbs-Resampling-Simulator vor. Eine Diskussion der Einschränkungen dieser beiden Simulationsverfahren sowohl hinsichtlich der Praktikabilität als auch hinsichtlich der asymptotischen Eigenschaften findet sich z.B. in Wilde (1999). Voraussetzung für das Erreichen der asymptotischen Effizienz von $\hat{\theta}_{MSS}$ in (24) bei Vernachlässigung dieser beiden Simulatoren ist, daß $\ddot{R}$ gegen unendlich wächst (vgl. McFadden, 1989, S. 1004).

Eine weitere von Hajivassiliou/McFadden (1998) vorgeschlagene Form der MSS-Schätzung ist die getrennte Simulation der Komponenten $f(Y_i|X_i;\theta)$ und $\partial f(Y_i|X_i;\theta)/\partial\theta$ von $S_i(\theta) = \partial ln f(Y_i|X_i;\theta)/\partial\theta$. Im allgemeinen Fall können diese Komponenten durch erwartungstreue Simulatoren

$$\tilde{f}(Y_i, X_i, V_i; \theta) \quad \text{bzw.} \quad \frac{\partial \widetilde{f(Y_i, X_i}, \ddot{V}_i; \theta)}{\partial\theta}$$

approximiert werden. Mit den aus den jeweiligen Verteilungen von V_i bzw. $\ddot{V}_i$ gezogenen V_{ir} $(i = 1, \ldots, N; r = 1, \ldots, R)$ und $\ddot{V}_{ir}$ $(i = 1, \ldots, N; r = 1, \ldots, \ddot{R})$ gelangt man zu folgendem speziellen MSS-Schätzer:

$$\hat{\theta}_{MSS} = argsolves \left[\sum_{i=1}^{N} \frac{\frac{1}{\ddot{R}} \sum_{r=1}^{\ddot{R}} \frac{\partial \widetilde{f(Y_i, X_i}, \ddot{V}_{ir}; \theta)}{\partial\theta}}{\frac{1}{R} \sum_{r=1}^{R} \tilde{f}(Y_i, X_i, V_{ir}; \theta)} = 0 \right] \tag{25}$$

Im MMPM ergibt sich mit einem Simulator $\frac{\partial \widetilde{P_{is}(\theta)}}{\partial\theta}$ für $\frac{\partial P_{is}(\theta)}{\partial\theta}$:

$$\hat{\theta}_{MSS} = argsolves \left[\sum_{i=1}^{N} \sum_{s\in S} Y_{is} \frac{\frac{\partial \widetilde{P_{is}(\theta)}}{\partial\theta}(\ddot{V}_{i1}^s, \ldots, \ddot{V}_{i\ddot{R}}^s)}{\tilde{P}_{is}(V_{i1}^s, \ldots, V_{iR}^s; \theta)} = 0 \right] \tag{26}$$

Allerdings werden durch diese Variante die $S_i(\theta)$ wegen der zugrunde liegenden nicht-linearen Transformation wiederum nicht erwartungstreu simuliert. Voraussetzung für die Konsistenz des MSS-Schätzers $\hat{\theta}_{MSS}$ in (25) bzw. (26) sowie der asymptotischen Normalverteilung von $\sqrt{N}(\hat{\theta}_{MSS} - \dot{\theta})$ ist, daß $N \to \infty$ und $R/\sqrt{N} \to \infty$. Die Anzahl $\ddot{R}$ an Simulationsreplikationen besitzt einen Einfluß auf die asymptotische Effizienz dieses MSS-Schätzers (vgl. Hajivassiliou/Ruud, 1994, S. 2432 f). Vorteilhaft ist es, wenn $V_{ir} = \ddot{V}_{ir}$ und damit $R = \ddot{R}$. Mit

$$\frac{\partial \widetilde{f(Y_i, X_i}, V_i; \theta)}{\partial\theta} = \frac{\partial \tilde{f}(Y_i, X_i, V_i; \theta)}{\partial\theta}$$

erhält man dann wieder den MSS-Schätzer in (22) sowie (23) und damit den SMLM-Schätzer in (8) bzw. (9).

7 Praktische Aspekte

Die asymptotischen Eigenschaften der SMLM, der SGMM sowie der MSS wurden in der Vergangenheit in der Literatur ausführlich analysiert. Hinsichtlich der Asymptotik sind auch die wesentlichen Vorzüge der SGMM sowie der MSS gegenüber der SMLM zu sehen (vgl. oben).

Bei dem Einsatz in empirischen Arbeiten ist für die Wahl eines bestimmten Schätzverfahrens aber insbesondere die Praktikabilität sowie das Verhalten bei beschränktem Beobachtungsumfang N und bei beschränkter Anzahl R an Simulationsreplikationen entscheidend.

Im Hinblick auf die Praktikabilität besitzt die SMLM insofern Vorteile, als für die Implementierung die üblichen Maximum-Likelihood-Programme verwendet werden können, lediglich ergänzt durch die Simulation der einzelnen Auswahlwahrscheinlichkeiten des MMPM in der Likelihoodfunktion. Insbesondere steht die SMLM in Verbindung z.B. mit dem GHK-Simulator im Rahmen der Programmpakete LIMDEP und GAUSSX direkt zur Verfügung. Bei der SGMM (und vor allem bei der modifizierten SGMM nach Keane) sowie bei verschiedenen Varianten der MSS ist dagegen eine viel speziellere Software notwendig.

Die wesentliche Einschränkung verschiedener Versionen der SGMM und der MSS ist jedoch die numerische Instabilität. Wiederholt tauchten bei eigenen Versuchen sowie in der Literatur (vgl. z.B. Mühleisen, 1994, Geweke u.a., 1994, Lee, 1995) bei der SGMM- und MSS-Schätzung Konvergenz- und Ausreißerprobleme auf. Im Hinblick auf die SGMM läßt sich dies damit erklären, daß die Minimierung der quadratischen Form numerisch schwierig ist. Zudem reagieren SGMM-Schätzungen häufig sehr sensibel auf verschiedene Startwerte für den iterativen Minimierungsprozeß. Bei der MSS steht man insbesondere bei der erwartungstreuen Simulation der Scores vor großen Problemen (vgl. sechsten Abschnitt). Dagegen werden in der Literatur bei der Maximierung der simulierten Likelihoodfunktion keinerlei derartige Schwierigkeiten beschrieben. Eigene Untersuchungen bestätigen die Stabilität von SMLM-Schätzungen.

Darüber hinaus ist zu berücksichtigen, daß die Güte der SGMM-Schätzung sehr stark von der Ausgestaltung der Instrumentenmatrix abhängt. So simulieren z.B. Geweke u.a. (1994) bei der SGMM-Schätzung die Instrumente zum Teil mit Hilfe der SMLM-Schätzer und verwenden diese auch als Startwerte des Suchalgorithmus. Damit liegt de facto ein zweistufiges Schätzverfahren vor, wodurch der direkte Vergleich dieses Schätzers mit der SMLM nicht legitim ist. Zudem dokumentiert Stern (1999) in seinen Monte-Carlo-Studien eine klare Überlegenheit der SMLM gegenüber der SGMM hinsichtlich der durchschnittlichen mittleren quadratischen Fehler in einperiodigen Sechs-Alternativen-Probitmodellen.

Falls man sich schließlich vergegenwärtigt, daß bei der SMLM im Gegensatz zur SGMM und MSS die Überprüfung statistischer Hypothesen (vgl. Lee, 1997b, 1999) sehr einfach ist, empfiehlt sich nach der Abwägung aller Vor- und Nachteile im Hinblick auf empirische Studien für die Parameterschätzung im MMPM die SMLM (insbesondere in Verbindung mit dem GHK-Simulator). Allerdings scheint die systematische Analyse der SMLM in Probitmodellen bei beschränktem Beobachtungsumfang N und bei beschränkter Anzahl R an Simulationsreplikationen in der Literatur (vgl. z.B. Börsch-Supan/Hajivassiliou, 1993, Geweke u.a., 1994, 1997, Keane, 1994, Lee, 1995, 1997a, Kaltenborn, 1997, Hyslop, 1999, Stern, 1999, Eymann/Ziegler, 2000) bei weitem noch nicht abgeschlossen. Hierin liegt auch im Hinblick auf die Untersuchung simulierter Testverfahren noch Forschungsbedarf.

Literatur

Amemiya T.(1985): Advanced Econometrics, Cambridge.

Asea P.K./Turnovsky S.J. (1998): Capital Income Taxation and Risk-Taking in a

Small Open Economy. *Journal of Public Economics*, 68, S. 55-90.

Bertschek I./Lechner M. (1998): Convenient Estimators for the Panel Probit Model. *Journal of Econometrics*, 87, S. 329-371.

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, S. 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, S. 347-368.

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, Chicago*, S. 79-104.

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, S. 477-498.

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

Eymann A./Ziegler A. (2000): Simulierte Maximum-Likelihood-Schätzung in Mehrperioden-Mehralternativen-Probitmodellen. *Unveröffentlichtes Manuskript, Universität Mannheim.*

Fomby T.B./Hill R.C./Johnson S.R. (1984): Advanced Econometric Methods, New York u.a.

Geweke J./Keane M./Runkle D. (1994): Alternative Computational Approaches to Inference in the Multinomial Probit Model. *The Review of Economics and Statistics*, Vol. LXXVI, No. 4, S. 609-632.

Geweke J./Keane M./Runkle D. (1997): Statistical Inference in the Multinomial Multiperiod Probit Model. *Journal of Econometrics*, 80, S. 125-165.

Gouriéroux C./Monfort A. (1991): Simulation Based Inference in Models with Heterogeneity. *Annales d'économie et de statistique*, No. 20/21, S. 69-107.

Gouriéroux C./Monfort A. (1993): Simulation-Based Inference. A Survey with Special Reference to Panel Data Models. *Journal of Econometrics*, 59, S. 5-33.

Gouriéroux C./Monfort A. (1995a): Statistics and Econometric Models, Vol. 1. General Concepts, Estimation, Prediction, and Algorithms, Cambridge.

Gouriéroux C./Monfort A. (1995b): Statistics and Econometric Models, Vol. 2. Testing, Confidence Regions, Model Selection, and Asymptotic Theory, Cambridge.

Gouriéroux C./Monfort A. (1996): Simulation-Based Econometric Methods, Oxford.

Hajivassiliou V.A. (1993): Simulation Estimation Methods for Limited Dependent

Variable Models. In: *Maddala G. S. u.a. (eds.): Handbook of Statistics, Vol. 11, Econometrics*, S. 519-543.

Hajivassiliou V.A. (1999): Some Practical Issues in Maximum Simulated Likelihood. Erscheint in: *Mariano R. u.a.: Simulation-Based Inference in Econometrics: Methods and Applications, Cambridge.*

Hajivassiliou V.A./McFadden D. (1998): The Method of Simulated Scores for the Estimation of LDV Models. *Econometrica*, Vol. 66, No. 4, S. 863-896.

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*, S. 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, S. 85-134.

Hansen L.P. (1982): Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, Vol. 50, No. 4, S. 1029-1054.

Hausman J.A./Wise D.A. (1978): A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. *Econometrica*, Vol. 46, No. 2, S. 403-426.

Hyslop D.R. (1999): State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. *Econometrica*, Vol. 67, No. 6, S. 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.*

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. (1993): Simulation Estimation for Panel Data Models with Limited Dependent Variables. In: *Maddala G.S. u.a. (eds.): Handbook of Statistics, Vol. 11, Econometrics*, S. 545-571.

Keane M. (1994): A Computationally Practical Simulation Estimator for Panel Data. *Econometrica*, Vol. 62, No. 1, S. 95-116.

Lechner M./Breitung J. (1996): Some GMM Estimation Methods and Specification Tests for Nonlinear Models. In: *Matyas L./Sevestre P. (eds.): The Econometrics of Panel Data. A Handbook of the Theory with Applications, 2nd rev. ed., Dordrecht/Boston/London*, S. 583-612.

Lee L.-F. (1992): On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Response Models. *Econometric Theory*, 8, S. 518-552.

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

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

Lee L.-F. (1997b): Some Common Structures of Simulated Specification Tests in Multinormal Discrete and Limited Dependent Variables Models. *Working Paper No. 97-4, The Hong Kong University of Science and Technology, Department of Economics.*

Lee L.-F. (1999): Statistical Inference with Simulated Likelihood Functions. *Econometric Theory*, 15, S. 337-360.

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, Cambridge/London*, S. 305-319.

McFadden D. (1989): A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, Vol. 57, No. 5, S. 995-1026.

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

Newey W.K. (1990): Efficient Instrumental Variables Estimation of Nonlinear Models. *Econometrica*, Vol. 58, No. 4, S. 809-837.

Newey W.K. (1993): Efficient Estimation of Models with Conditional Moment Restrictions. In: *Maddala G.S. u.a. (eds.): Handbook of Statistics, Vol. 11, Econometrics*, S. 419-454.

Ronning G. (1991): Mikroökonometrie, Berlin u.a.

Ruud P. (1998): An Introduction to Classical Econometric Theory.

Stern S. (1999): Simulation Based Inference in Econometrics: Motivation and Methods. Erscheint in: *Mariano R. u.a.: 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, S. 281-307.

Weeks M. (1997): The Multinomial Probit Model Revisited: A Discussion of Parameter Estimability, Identification and Specification Testing. *Journal of Economic Surveys*, Vol. 11, No. 3, S. 297-320.

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