„Ökonometrische Analyse von aggregierten Tendenzdaten aus Panelerhebungen“*

Von Gerd Ronning

Mit Tendenzfragen, die heute in vielen Gebieten der Markt-, Meinungs- und Wirtschaftsforschung verwendet werden, läßt sich in Umfragen die zeitliche Veränderung von bestimmten Merkmalen (z. B. Preiserwartungen, Produktionspläne, Kapazitätsauslastung) sehr einfach erfassen. Deshalb ist diese — zeitsparende — Art der Erhebung vor allem für oftmals wiederholte Befragungen in Panels von Interesse. In diesem Artikel wird ein Ansatz diskutiert, der die qualitative Information aus einer Tendenzfrage durch beliebige Einflußgrößen erklärt. Dabei wird unterstellt, daß nur aggregierte Daten zur Verfügung stehen.

1. Einleitung

1.1 Ziel der Arbeit

Der ursprüngliche Anlaß für die Beschäftigung mit Tendenzdaten war die Tatsache, daß diese Daten — unter anderem — bei der empirischen Erfassung von (Unternehmer-)Erwartungen bezüglich ökonomischer Variablen eine Rolle spielen. Da die Häufigkeit von Tendenzbefragungen aber auch in vielen anderen Bereichen (Abfrage des Ist-Zustands, Befragung von Haushalten) wächst, können die folgenden Ausführungen für alle Bereiche, in denen Tendenzdaten aus Panels gewonnen werden, als relevant angesehen werden, obwohl sich die Darstellung weitgehend am Konjunkturtest (*KT*) des Münchner IFO-Instituts orientiert.

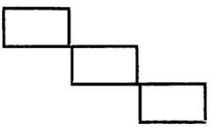
Die Attraktivität der Tendenzbefragung liegt darin, daß die Beantwortung sehr einfach ist und deshalb wenig Zeit erfordert. Zwei typische Tendenzfragen aus dem IFO-*KT* zeigt Abbildung 1/1. Nachteilig wirkt sich der Informationsverlust aus, der dadurch ensteht, daß nur die Tendenz („−“, „=“, „+“) abgefragt wird. Dies gilt insbesondere bei trendhaften Daten wie z. B. Preisen.

---

* Revidierte (und in Abschnitt 2.4 ergänzte) Fassung eines Vortrags vor dem Ökonometrischen Ausschuß des Vereins für Socialpolitik in Bad Honnef, März 1980. Bei der Durchführung der Simulationsexperimente hat Herr Dipl. Volksw. *Arnold Holz* mitgewirkt. Diese Arbeit ist Teil des Projekts 20/76, das von der Universität Konstanz finanziell unterstützt wird.

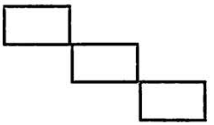
Unsere Produktionstätigkeit*) bezüglich XY wird voraussichtlich im Laufe der nächsten 3 Monate in konjunktureller Hinsicht — also unter Ausschaltung rein saisonaler Schwankungen —

steigen
etwa gleich bleiben
abnehmen



Unsere Inlandsverkaufspreise (Nettopreise) für XY werden — unter Berücksichtigung von Konditionsveränderungen — voraussichtlich im Laufe der nächsten 3 Monate

steigen
etwa gleich bleiben
abnehmen



*) Veränderungen, die lediglich auf übliche Betriebsferien bzw. Urlaube und regelmäßig wiederkehrende Reparaturen sowie eine unterschiedliche Monatslänge zurückgehen, sind nicht zu berücksichtigen.

Abbildung 1/1 — Tendenzfragen (IFO-Konjunkturtest)

Ökonometrische Modelle, in denen Tendenzdaten verwendet werden, müssen dem speziellen Charakter dieser Daten Rechnung tragen. Dabei ist auch zu unterscheiden, ob Mikro- oder Makrodaten zur Verfügung stehen. Die neueste ökonometrische Literatur hat vorwiegend Modelle für Mikrodaten betrachtet, obwohl nach wie vor aus Geheimhaltungs- und/oder Kostengründen sehr häufig nur Makrodaten zur Verfügung stehen und entsprechende Modellformulierungen erfordern. In dieser Arbeit wird das logistische Modell zur Erklärung von *aggregierten* Tendenzdaten benutzt. Ziel der Arbeit ist es, Vor- und Nachteile dieses Erklärungsansatzes darzustellen.

1.2 Allgemeine Bemerkungen

In *Produzenten*panels können neben den üblichen Problemen der Panelanalyse auch strategisch motivierte absichtliche Fehlmeldungen eine Rolle spielen, insbesondere bei oligopolistischen Marktstrukturen[1]. Ferner tendieren die Panelmitglieder mit der Zeit dazu, sich stark an Angaben für frühere Zeitpunkte zu orientieren. Für den IFO-*KT* siehe dazu *Vogler* (1978). Bei saisonal beeinflußten Tatbeständen (z. B. Erwartungen über Verkaufspreise) kann dies ein durchaus sinnvolles Antwortverhalten sein. Siehe dazu Abschnitt 2.3. Diese Bemerkungen sollen andeuten, daß bei Daten aus Produzentenpanels die Annahme stochastischer Unabhängigkeit für den datenerzeugenden Prozeß sowohl im Querschnitt als auch im Längsschnitt verletzt sein dürfte.

[1] Dies kann für die Frage, ob Unternehmermeldungen gewichtet werden sollen, von Bedeutung sein. Siehe dazu die folgenden Ausführungen.

Tendenzfragen beziehen sich auf die *Veränderung* einer quantitativen Variablen, wobei übrigens offenbleibt, ob der Befragte relative oder absolute Änderungen meint. Die Antwort wird durch eine polytome (im IFO-*KT* trichotome) Variable beschrieben, wobei die einzelnen Antwortkategorien („−", „=", „+") in einer natürlichen *Rangordnung* stehen.

Da jedes Unternehmen unabhängig von seinem Marktanteil nur „eine Stimme" hat, ist es sinnvoll, die *Einzelmeldungen* zu *gewichten,* bevor sie aggregiert werden. Allerdings ist die Gewichtung von Tendenzdaten wegen des Informationsverlustes problematischer als jene von quantitativen Erhebungsergebnissen. Darüber hinaus sind extreme Marktkonstellationen denkbar, in denen das *un*gewichtete Aggregat ein sinnvolleres Maß ist. Statistisch gesehen führt die Gewichtung zu einer veränderten Stichprobenverteilung. Darauf wird jedoch im folgenden nicht eingegangen[2].

Modelle mit *diskreten* abhängigen Variablen, die zur Analyse von Tendenzdaten aus Panelerhebungen geeignet wären, sind bisher kaum entwickelt worden. Dieser Eindruck wird bestätigt durch die Lektüre eines Konferenzbandes mit dem Titel "The Econometrics of Panel Data"[3], in dem sich nur ein einziger Artikel (*Heckman* 1978) mit diskreten Paneldaten beschäftigt, der allerdings von *Nerlove* im Vorwort als der „vielleicht bedeutsamste methodologische Aufsatz der Konferenz" bezeichnet wird (INSEE 1978, S. 12).

Es ist müßig darüber zu streiten, ob man *Makrodaten*[4] oder *Mikrodaten* in Modellen benutzen soll. Wegen des Informationsverlustes bei der Aggregation (gleich ob über Personen, Produkte, Zeit o. ä.) sollte man Mikrodaten verwenden, wenn man sie hat. Aber man hat sie eben oft *nicht.* Bei linearen Modellen kann die Aggregation häufig modelliert werden (z. B. *Schneeweiss* 1978, Kap. 3.3.3). Für den Fall *aggregierter Tendenzdaten* wird in Abschnitt 2 demonstriert, daß eine Spezifikation und Schätzung mittels des logistischen Modells sinnvoll sein kann. Probleme ergeben sich im Fall aggregierter Tendenzdaten dann, wenn *mehrere* diskrete Variable (z. B. Ergebnisse für verschiedene Fragen aus dem *KT*) gemeinsam analysiert werden sollen[5] oder die endogene

---

[2] Die statistische Analyse wird weiter dadurch erschwert, daß die Anzahl der Meldungen im Panel über die Zeit hin nicht konstant ist. Auch dieser Aspekt bleibt im folgenden unberücksichtigt.

[3] Genaue Quelle im Literaturverzeichnis unter INSEE (1978).

[4] Ich verwende die beiden Begriffe „aggregierte Daten" und „Makrodaten" synonym.

[5] Siehe *Nerlove* und *Press* (1978) und *Bishop, Fienberg* und *Holland* (1975, Kap. 10.4.4), ferner im Zusammenhang mit den Ifo-Tendenzdaten *König* und *Nerlove* (1980) sowie *König, Nerlove* und *Oudiz* (1979).

Variable um eine oder mehrere Perioden verzögert als Einflußgröße spezifiziert werden soll.

1.3 Terminologie und Symbolik

Um im folgenden eine einheitliche Terminologie und Symbolik benutzen zu können, sei diese hier insgesamt vorgestellt. Dabei beschränke ich mich auf den *univariaten* Fall, d. h. ich betrachte Modelle mit nur einer (diskreten) abhängigen Variablen, z. B. Preiserwartungen oder Produktionsplänen aus Abbildung 1/1. Mit

$$p_j(t) \ , \qquad j = 1, \ldots, r \ ; \quad t = 1, \ldots, T$$

wird die Wahrscheinlichkeit bezeichnet, daß bei der Befragung im Zeitpunkt t die j-te Antwortkategorie gewählt wird. Für trichotome Tendenzdaten gibt es $r = 3$ Kategorien, nämlich „$-$“ ($j = 1$), „$=$“ ($j = 2$) und „$+$“ ($j = 3$) in offenbarer Symbolik für die Richtung der Veränderung. Entsprechend sei $m_j(t)$ die Anzahl der Untersuchungseinheiten, die im Zeitpunkt t die j-te Kategorie wählen. Die Gesamtzahl Fälle pro Zeitpunkt ist dann durch

$$n(t) = \sum_{j=1}^{r} m_j(t) \ , \quad t = 1, \ldots, T$$

gegeben. Ferner seien $y_j(t)$ die entsprechenden relativen Anteile in einem bestimmten Zeitpunkt, d. h.

$$y_j(t) = m_j(t) / n(t) \ .$$

Mit

$$x_{tk} \ , \quad k = 1, \ldots, K \ ; \quad t = 1, \ldots, T$$

wird der Wert der k-ten Einflußvariablen im Zeitpunkt t bezeichnet, und

$$x'(t) = (x_{t1}, x_{t2}, \ldots, x_{tK})$$

bezeichnet den K-dimensionalen Vektor der Einflußgrößen in t. Ferner ist

$$\beta'_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{Kj})$$

ein K-dimensionaler Vektor für die j-te Kategorie und

$$\beta' = (\beta'_1, \beta'_2, \ldots, \beta'_{r-1})$$

ein $K(r-1)$-dimensionaler Vektor.

Wenn für eine bestimmte Untersuchungseinheit für jeden Zeitpunkt t ($t = 1, \ldots, T$) bezüglich der untersuchten Variablen (Tendenzfrage im *KT*) die Information über die gewählte Kategorie vorliegt, dann verfügt man über *Mikrodaten* für diese Untersuchungseinheit, insbesondere auch über (i, j)-Kombinationen in $t-1$ und t. Liegen dagegen für die betrachtete Variable nur Informationen in der Form $y_j(t)$, $j = 1, \ldots, r$, und $n(t)$ vor, so spreche ich von *aggregierten Tendenzdaten* (bezüglich der betrachteten Variablen).

2. Ein logistisches Modell zur Erklärung von aggregierten Tendenzdaten

2.1 Modellspezifikation

Im folgenden wird unterstellt, daß die Antwortwahrscheinlichkeiten $p_j(t)$ bezüglich einer bestimmten Tendenzfrage von K verschiedenen exogenen Variablen beeinflußt werden. Wir benutzen dafür das logistische Modell. Dann erhalten wir

$$p_j(t) = \frac{\exp(\beta_j' x(t))}{\sum_{i=1}^{r} \exp(\beta_i' x(t))}, \quad j = 1, \ldots, r, \tag{2-1}$$

wobei der Einflußgrößenvektor $x(t)$ als erstes Element jeweils eine 1 enthalten soll, d. h. $x_{t1} \equiv 1$ für alle t. Im Fall $K = 2$, in dem nur eine exogene Variable auf die Antwortwahrscheinlichkeiten einwirkt, erhalten wir aus (2-1):

$$p_j(t) = \frac{\exp(\beta_{1j} + \beta_{2j} x_{t2})}{\sum_{i=1}^{r} \exp(\beta_{1i} + \beta_{2i} x_{t2})} =: p_j(t, x_{t2}) . \tag{2-2}$$

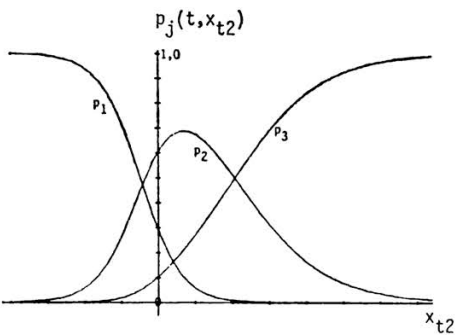


Abbildung 2/1 — Antwortwahrscheinlichkeiten im logistischen Modell

Abbildung 2/1 stellt die Antwortwahrscheinlichkeiten für das spezielle Modell (2-2) in Abhängigkeit von x_{t2} graphisch dar[6]. Dabei erfüllen die β_{2j}-Koeffizienten die Restriktionen

$$\beta_{21} < \beta_{22} < \beta_{23} \; . \tag{2-3}$$

Man erkennt aus der Abbildung 2/1, daß die Antwortwahrscheinlichkeiten p_1 für „−" und p_3 für „+" monoton sinken bzw. wachsen, wenn x_{t2} wächst, während die Antwortwahrscheinlichkeit p_2 für „=" erst ansteigt und dann wieder abfällt. Es läßt sich leicht formal zeigen, daß dies unterschiedliche Verhalten der Antwortwahrscheinlichkeiten für „extreme" und „mittlere" Antwortkategorien genau dann vorliegt, wenn die Restriktionen (2-3) erfüllt sind. Dies ermöglicht es dem Ökonometriker, im Fall der Verwendung von (2-2) direkt zu überprüfen, ob seine Schätzwerte für β_{2j}, $j = 1, \ldots, r$, sinnvoll sind. Dies wird in Abschnitt 2.4 anhand eines Beispiels illustriert.

Für den Fall mehrerer Einflußvariablen ($K > 2$) ist eine entsprechende Analyse nicht möglich. Man könnte in diesem Fall logistische Modelle verwenden, in denen die Ordnung der Antwortkategorien explizite berücksichtigt ist. Siehe dazu *Bock* (1975) und *Amemiya* (1975, 1976). Dieser Aspekt, der offensichtlich auch für Tendenzdaten relevant ist, soll in einer geplanten Arbeit weiterverfolgt werden.

2.2 Logitanalyse für aggregierte Tendenzdaten

Wir betrachten im folgenden das Modell (2-1), das die Antworthäufigkeiten in Form aggregierter Tendenzdaten in Abhängigkeit von K Einflußgrößen erklären soll. Dabei wird vorausgesetzt, daß für T aufeinanderfolgende Zeitpunkte sowohl der Vektor $x(t)$ als auch Information von der Form $m_j(t)$, $j = 1, \ldots, r-1$, und $n(t)$ vorhanden ist. Damit entspricht unser Modell der üblichen Vorstellung in der Ökonometrie, daß eine bestimmte Struktur (dargestellt durch die β_j-Vektoren in (2-1)) über T Perioden konstant existierte und ein datenerzeugender Prozeß in Abhängigkeit von dieser Struktur die Beobachtungswerte generierte. Dabei wird in beiden zu besprechenden Schätzmethoden angenommen, daß sowohl die Beobachtungen eines Zeitpunkts t als auch Beobachtungen für verschiedene Zeitpunkte stochastisch voneinander unabhängig sind. Es wird unterstellt, daß die $n(t)$ Beobachtungen in den T verschiedenen Zeitpunkten jeweils durch eine Multinomialverteilung mit Parametern $p_j(t)$, $j = 1, \ldots, r-1$ und $n(t)$ erzeugt wurden. Man nennt dies Modell auch „Mehrfach-Multinomial-Stichprobe" (product multinomial sampling). Siehe *Bishop, Fienberg* und *Holland*

[6] Die gewählte Parameterkonstellation wird in der Simulationsstudie benutzt. Siehe Beispiel II in Abschnitt 2.4.

(1975, 63). Es sollte hier nochmals darauf hingewiesen werden, daß dieses Stichprobenmodell in mehrfacher Hinsicht verletzt sein kann. Einmal sind die T verschiedenen Stichproben im Panel vermutlich nicht voneinander unabhängig. Ferner ist $n(t)$ im allgemeinen nicht fix, weil die Anzahl der Meldungen schwankt[7] und schließlich werden die Meldungen oftmals noch gewichtet.

Unter der Annahme der Mehrfach-Multinomial-Stichprobe ergibt sich bei Vorgabe der Wahrscheinlichkeiten (2-1) die Likelihoodfunktion

$$(2\text{-}4) \qquad \mathfrak{L}\,(m_j(t);\ j=1,\ \ldots,\ r;\ t=1,\ \ldots,\ T \mid n(t), x(t), \beta) = \prod_{t=1}^{T} n(t)! \prod_{j=1}^{r} \frac{p_j(t)^{m_j(t)}}{m_j(t)!}$$

a) Normierung

Das Modell (2-1) ist nur dann eindeutig parametrisiert, wenn eine Normierung eingeführt wird[8]. Üblicherweise werden — ähnlich wie bei linearen Modellen — die folgenden Normierungen verwendet:

$$(2\text{-}5) \qquad \sum_{j=1}^{r} \beta_{kj} = 0\ , \quad k = 1,\ \ldots,\ K \quad (\text{„Symmetrische Normierung“})$$

$$(2\text{-}6) \qquad \beta_{kr} = 0\ , \quad k = 1,\ \ldots,\ K \quad (\text{„Restkategorie-Normierung“})$$

Dabei wird (2-5) vor allem in der *ML*-Schätzung, (2-6) dagegen in der *BTM*-Schätzung verwendet (siehe Unterabschnitt b). Da die verschiedenen Normierungen zu verschiedenen Schätzwerten für den Koeffizientenvektor β und dessen Kovarianzmatrix (dagegen natürlich nicht zu verschiedenen geschätzten Wahrscheinlichkeiten) führen, ist es wichtig, Formeln für den Zusammenhang zwischen den beiden Parametrisierungen zu kennen. Wenn man die Parameter unter der Restkategorie-Normierung mit β^{*}_{kj} und die unter der symmetrischen Normierung mit β^{**}_{kj} bezeichnet, dann läßt sich zeigen:

$$(2\text{-}7) \qquad \beta^{*}_{kj} = \beta^{**}_{kj} + \sum_{i=1}^{r-1} \beta^{**}_{ki}$$

$$(2\text{-}8) \qquad \beta^{**}_{kj} = \beta^{*}_{kj} - \frac{1}{r} \sum_{i=1}^{r-1} \beta^{*}_{ki}$$

Will man beispielsweise Schätzungen unter der Restkategorie-Normierung (2-6) in Werte unter der symmetrischen Normierung (2-5) um-

---

[7] Siehe auch Fußnote 2.

[8] Wenn man den Bruch in (2-1) mit $\exp(\gamma' x(t))$, γ ein beliebiger K-dimensionaler Vektor, erweitert, dann erhält man das logistische Modell mit Parametervektoren $(\gamma + \beta_j)$ anstelle von β_j.

wandeln, so setzt man die Schätzwerte an die Stelle von β^*_{kj} in (2-8) und berechnet β^{**}_{kj}. Die dazugehörenden Varianzen berechnen sich dann wie folgt:

$$\sigma^{**}_{jj}(k) = \frac{1}{r^2}\left((r-1)^2\,\sigma^*_{jj}(k) + \sum_{\substack{i=1\\ i\neq j}}^{r-1} \sigma^*_{ii}(k) - 2(r-1)\sum_{\substack{i=1\\ i\neq j}}^{r-1} \sigma^*_{ij}(k)\right), \quad j = 1, \ldots, r\,; \quad k = 1, \ldots, K \tag{2-9}$$

Dabei ist $\sigma_{ij}(k)$ die Kovarianz für Schätzer $\tilde{\beta}_{kj}$ und $\tilde{\beta}_{kj}$ für die betreffenden Koeffizienten, und die Sterne bezeichnen die jeweilige Normierung.

b) Zwei alternative Schätzmethoden

Nach wie vor ist meiner Ansicht nach nicht eindeutig geklärt, ob die Maximum-Likelihood-Schätzmethode (*ML*-Methode) der zuerst von *Berkson* vorgeschlagenen und von *Theil* (1967, 1970) in den Sozialwissenschaften populär gemachten Schätzmethode (Abkürzung: *BTM* = Berkson-Theil-Methode) im Zusammenhang mit dem logistischen Modell vorzuziehen ist[9]. Beide Methoden seien im folgenden kurz vorgestellt. Jedesmal taucht die $[(r-1)K \times (r-1)K)]$-Matrix

$$H = \sum_{t=1}^{T} \underset{[(r-1)\times(r-1)]}{V_t} \otimes \underset{[K\times K]}{x(t)\,x'(t)} \tag{2-10}$$

auf, wobei die Elemente $v_{ij}(t)$ in V_t folgende Form besitzen:

$$v_{ij}(t) = \left\{\begin{array}{ll} n(t)\,\tilde{p}_i(t)\,(1-\tilde{p}_i(t)), & i=j \\ -\,n(t)\,\tilde{p}_i(t)\,\tilde{p}_j(t), & i\neq j \end{array}\right\} i,j = 1, \ldots, r-1, \tag{2-11}$$

und $\tilde{p}_i(t)$ bestimmte Schätzwerte für die unbekannten Wahrscheinlichkeiten in (2-1) sind. Man erkennt, daß V_t die Kovarianzmatrix einer Multinomialverteilung mit $n(t)$ Versuchen ist (*Bishop, Fienberg* und *Holland* 1975, S. 442), die stets positiv definit ist, sofern alle r Wahrscheinlichkeiten positiv sind und $n(t) > 0$ gilt. Macht man ferner die in der Ökonometrie übliche Annahme, daß für K verschiedene Zeitpunkte linear voneinander unabhängige Vektoren $x(t)$ existieren, dann ist die in (2-10) gegebene Matrix H positiv definit. Dies ist insofern bedeutsam, als H^{-1} in beiden Schätzmethoden als (asymptotische) Kovarianzmatrix auftaucht. Zwei alternative Beweise der Definitheit von

[9] Eine allgemeine Formulierung beider Methoden für den polytomen Fall ist bei *Amemiya* (1976) zu finden. Für den dichotomen Fall werden empirische Schätzungen bei *Nerlove* und *Press* (1973) und Simulationsergebnisse bei *Domencich* und *McFadden* (1975) präsentiert.

H liegen z. B. von *Bock* (1975, 525) und *Dhrymes* (1978, 347 ff.) vor. Ferner könnte man ausnutzen, daß

$$H = (X \otimes I)' \left(V_t \overset{T}{\underset{t=1}{\otimes}} I\right) (X \otimes I) \tag{2-12}$$

gilt; dabei ist $X' = (x(1), x(2), \ldots, x(T))$ die $[K \times T]$-Einflußgrößenmatrix für alle T Zeitpunkte und $\overset{T}{\underset{t=1}{\otimes}}$ bezeichnet das in *Lee, Judge* und *Zellner* (1977, 76) eingeführte „zeitlich geordnete Kronecker-Produkt".

Die beiden Schätzmethoden lassen sich nun wie folgt beschreiben: Der *BTM-Schätzer* ist *bei Vorgabe der Restkategorie-Normierung* (2-6) durch

$$\hat{\beta} = \hat{H}^{-1} (X \otimes I_{r-1})' \left(\hat{V}_t \overset{T}{\underset{t=1}{\otimes}} I_T\right) L \tag{2-13}$$

mit asymptotischer Kovarianzmatrix $\hat{H}^{-1}$ gegeben. Dabei ist $L' = (L_1(1), \ldots, L_1(T), L_2(1), \ldots, L_2(T), \ldots, L_{r-1}(1), \ldots, L_{r-1}(T))$ der Vektor der empirischen „Logits"

$$L_j(t) = \log (\hat{p}_j(t) / \hat{p}_r(t)) \tag{2-14}$$

und die geschätzten Wahrscheinlichkeiten in den Matrizen $\hat{H}$ und $\hat{V}_t$ sowie in (2-14) werden dabei üblicherweise durch

$$\hat{p}_j(t) = m_j(t) / n(t) \tag{2-15}$$

bestimmt. Für kleines $n(t)$ ist diese Schätzung jedoch unbefriedigend. Außerdem ist dann der Fall $m_j(t) = 0$ besonders häufig, für den der *BTM*-Schätzer (2-13) bei Verwendung von (2-15) nicht definiert ist. Siehe (2-14). Da laut Modellannahme alle Wahrscheinlichkeiten $p_j(t)$ positiv sein sollen, ist folgende Schätzung eventuell vorzuziehen:

$$\hat{p}_j(t) = \frac{n(t)}{n(t)+1}\, y_j(t) + \frac{1}{n(t)+1}\,\frac{1}{r} \quad (\text{„}BMT1\text{"}) \; . \tag{2-16a}$$

Sie kann als (datenunabhängige) pseudo-bayessche Schätzung für $p_j(t)$ angesehen werden: Es wird a priori Gleichwahrscheinlichkeit der r Kategorien auf der Informationsgrundlage einer einzigen Beobachtung unterstellt. Alternativ wird in dieser revidierten Fassung eine *BTM*-Schätzung vorgeschlagen, die an die Stelle von $1/r$ in (2-16a) eine empirische Schätzung von der Form $\sum_t m_j(t) / \sum_t n(t)$ setzt:

$$\hat{p}_j(t) = \frac{n(t)}{n(t)+1}\, y_j(t) + \frac{1}{n(t)+1}\,\frac{\sum_{s=1}^{T} m_j(s)}{\sum_{s=1}^{T} n(s)} \quad (\text{„}BTM2\text{"}) \tag{2-16b}$$

Diese Wahrscheinlichkeitsabschätzung kann allerdings — im Gegensatz zu (2-16a) — den Wert Null ergeben, sofern für bestimmtes j alle $m_j(t)$ den Wert Null annehmen. Für andere Möglichkeiten der Abschätzung siehe *Bishop, Fienberg* und *Holland* (1975, 408). Je größer $n(t)$ wird, desto weniger unterscheiden sich die drei letzten Formeln.

Der *ML*-Schätzer für die unbekannten Koeffizienten in (2-1) ergibt sich aus der ersten Ableitung der logarithmierten Likelihoodfunktion (2-4). Unter Beachtung der *symmetrischen Normierung* erhält man

$$\frac{\delta \log \mathfrak{L}}{\delta \beta_j} = \sum_{t=1}^{T} [(m_j(t) - m_r(t)) - n(t)(p_j(t) - p_r(t))]\, x(t) \;, \quad j = 1, \ldots, r-1 \tag{2-17}$$

Wenn wir diesen Vektor der ersten Ableitungen für ein bestimmtes j gleich dem Nullvektor setzen, dann gilt

$$\sum_{t=1}^{T} [m_r(t) - n(t)\,\check{p}_r(t)]\, x(t) = 0 \;, \tag{2-18}$$

so daß wir die Extremwertbedingungen wie folgt schreiben können:

$$\sum_{t=1}^{T} [m_j(t) - n(t)\,\check{p}_j(t)]\, x(t) = 0 \;, \quad j = 1, \ldots, r-1 \tag{2-19}$$

Die Hessesche Matrix der zweiten Ableitungen der logarithmierten Likelihoodfunktion hat folgendes Aussehen:

$$\frac{\delta^2 \log \mathfrak{L}}{\delta\beta\, \delta\beta'} = -\sum_{t=1}^{T} Q_t \otimes x(t)\, x'(t) \tag{2-20}$$

Dabei sind die Elemente der $[(r-1) \times (r-1)]$-Matrix Q_t wie folgt[10]:

$$q_{ij}(t) = \begin{cases} n(t)\,[p_i(t)(1 - p_i(t) + p_r(t)(1 - p_r(t)) + 2\,p_i(t)\,p_r(t)] & , i = j \\ n(t)\,[-p_i(t)\,p_j(t) + p_r(t)(1 - p_r(t)) + p_r(t)(p_i(t) + p_j(t))] & , i \neq j \end{cases} \tag{2-21}$$

Wenn eine Lösung bezüglich β für das Gleichungssystem (2-19) existiert, dann ist diese Lösung der *ML*-Schätzer und wir erhalten im Punkte der Lösung von (2-19):

$$\frac{\delta^2 \log \mathfrak{L}}{\delta\beta\, \delta\beta'} = -\check{H} \tag{2-22}$$

mit H aus (2-10) und geschätzten Wahrscheinlichkeiten $\check{p}_j(t)$ auf der

[10] Die Matrix Q_t ist positiv definit, sofern alle Wahrscheinlichkeiten positiv sind und $n(t) > 0$ gilt. Den Beweis verdanke ich Wilhelm Forst, Mathematische Fakultät der Universität Konstanz.

Basis der *ML*-Schätzung für den Vektor β. Da die Hessesche Matrix (2-22) (mit $p_j(t)$ anstelle von $\check{p}_j(t)$) nicht-stochastisch ist, ergibt sich H^{-1} als asymptotische Kovarianzmatrix des *ML*-Schätzers.

c) *Ein Vergleich der beiden Schätzmethoden*

Beide Schätzer sind (für festes T und wachsendes $n(t)$) konsistent. Für *ML* siehe dazu z. B. *Dhrymes* (1978, 340). Für *BTM* folgt dies aus der Konsistenz der Schätzungen (2-15) und (2-16) bezüglich $p_j(t)$. Wie oben gezeigt, ist für beide Methoden die asymptotische Kovarianz durch H^{-1} *gegeben.* Siehe dazu auch *Amemiya* (1976), der die Matrix H mit $p_j(t)$ anstelle von $\tilde{p}_j(t)$ betrachtet.

Die *BTM*-Schätzung ist nur definiert, wenn — im Falle von (2-15) — keine Nullbeobachtungen für $m_j(t)$ vorliegen, d. h. es muß notwendigerweise $n(t) \geq r$ gelten. Ferner sind die Modifikationen in (2-16a) und (2-16b) nur dann sinnvoll, wenn $n(t)$ nicht zu klein ist. Im Fall von Mikrodaten kann dies nur durch „Gruppierung" erreicht werden. Die (negativen) Auswirkungen auf die Schätzungen im Falle von (2-15) sind bei *Theil* (1967, Kap. 3) ausführlich anhand von Simulationsergebnissen demonstriert worden. Weitere Ergebnisse dazu finden sich bei *Domencich* und *McFadden* (1975, Kap. 5.4 und 5.5). Für den Fall von aggregierten Daten, in denen die individuellen Reize nicht modelliert werden können, erweist sich andererseits diese Schätzmethode als attraktiv, sofern $n(t)$ nicht zu klein ist. Dies wird auch aus den in Abschnitt 2.4 unten dargestellten Simulationsergebnissen deutlich.

Auch die *ML*-Schätzung, die im übrigen numerisch-iterativ bestimmt werden muß[11], hat ihre Probleme bei kleinen Stichprobenumfängen, sofern Nullbeobachtungen vorliegen. Siehe *Nerlove* und *Press* (1973, 69). Man betrachte beispielsweise den Fall, daß die k-te Einflußvariable nur zwei Werte annimmt, $x_{tk} = c$ oder $x_{tk} = 0$, und für ein bestimmtes j das Ereignis $m_j(t) = 0$ stets dann eintritt, wenn $x_{tk} = c$ gilt. Für die k-te Zeile in (2-19) erhalten wir dann

$$-c \sum_t n(t) \check{p}_j(t) = 0 \ .$$

Diese Gleichung ist nur dann erfüllt, wenn für jedes t $\check{p}_j(t) = 0$ gilt, d. h. die Koeffizientenwerte im Vektor β (absolut) unendlich groß sind. Für die aggregierten IFO-Tendenzdaten findet sich dazu ein empirisches Beispiel in *Ronning* (1980): Die stark saisonal variierenden

[11] Dabei kann der Newton-Raphson-Algorithmus verwendet werden. Neuerdings wird auch die Modifikation von Davidon (*Domencich* und *McFadden* 1975, 121) oder der Fletcher-Powell-Algorithmus (*Nerlove* und *Press* 1978, 85) benutzt.

Preiserwartungen (siehe Abbildung 2/22) sollten durch beobachtete Preisveränderungen gemeinsam mit Saisonvariablen erklärt werden. Da jedoch z. B. jeweils im Dezember kein Unternehmen erwartete *sinkende* Preise meldete, ergab sich keine Konvergenz bei der iterativen Bestimmung der *ML*-Schätzung, während bei der *BTM*-Schätzung (unter Verwendung von (2-16a)) keine numerischen Probleme auftraten.

d) Güte der Anpassung

An die Stelle der Bestimmtheitsmaße treten bei qualitativen abhängigen Variablen Maße, die die Differenz zwischen beobachteten und geschätzten Häufigkeiten messen. Dabei verwendet man entweder das Likelihoodquotienten-Maß

$$\chi^2_{LR} = 2 \sum_{t=1}^{T} \sum_{j=1}^{r} m_j(t) \log (m_j(t)/(n(t)\,\tilde{p}_j(t))) \tag{2-23}$$

oder das Maß von Pearson

$$\chi^2_{PE} = \sum_{t=1}^{T} \sum_{j=1}^{r} (m_j(t) - n(t)\,\tilde{p}_j(t))^2/(n(t)\,\tilde{p}_j(t)) \tag{2-24}$$

(*Bock* 1975, 527). Vorausgesetzt die geschätzten (absoluten) Häufigkeiten $n(t)\,\tilde{p}_j(t)$ sind nicht zu klein[12] (umstrittene Faustregel: $n(t)\,\tilde{p}_j(t) \geq 5$), dann sind beide Maße asymptotisch χ^2-verteilt mit $(T - K)(r - 1)$ Freiheitsgraden. In der *ML*-Schätzung scheint (2-23) üblicher zu sein (z. B. *Haberman* 1978), während mir für einen Vergleich von *ML* und *BTM* das Maß (2-24) neutraler und deshalb angemessener zu sein scheint. In den empirischen Ergebnissen des Abschnitts 2.3 wird deshalb nur das Pearson-Maß verwendet. In den Simulationsexperimenten (Abschnitt 2.4) wird dann für beide Maße überprüft, ob die oben behauptete Verteilungseigenschaft für die betrachteten Stichprobenumfänge verwendet werden kann.

2.3 Empirische Ergebnisse

In einer früheren Arbeit (*Ronning* 1980) habe ich versucht, aggregierte Tendenzdaten aus dem IFO-*KT*, Produktgruppe „Werkzeugmaschinen spanabhebend", durch verschiedene Einflußgrößen zu erklären. Dabei wurde das Modell (2-1) zugrundegelegt. Zwei dieser Schätzungen sollen hier als Beispiele präsentiert werden. Sie beziehen sich auf die Preiserwartungen und Produktionspläne im Zeitraum 1973 bis

[12] Bei dieser Gelegenheit sei darauf hingewiesen, daß für die Verwendung der Normalverteilungsannahme für die Schätzer sogar $n(t)\,\tilde{p}_j(t) > 25$ gefordert wird. Siehe *Haberman* (1974, 144 - 146).

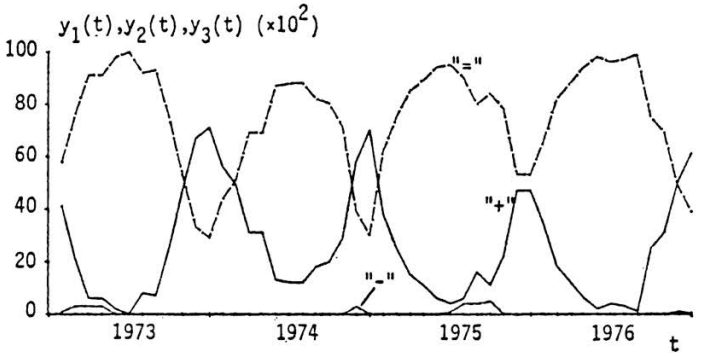
1977 ($T = 48$ Monate); die zugrundeliegenden Fragen sind in Abbildung 1/1 aufgeführt, die benutzten Daten werden in Abbildung 2/2, Teile a) und c) auf S. 194, dargestellt. Ferner ist in Teil b) der Verlauf der realisierten einmonatigen Wachstumsraten der Verkaufspreis dieser Produktgruppe zu sehen. Man beachte die ausgeprägte Saisonfigur bei erwarteten und realisierten Preisveränderungen. Aus der erwähnten Arbeit werden hier zwei Ergebnisse in den Tabellen 2.1 und 2.2 vorgestellt. In Tabelle 2.1 werden die Preiserwartungen in Abhängigkeit von beobachteten 3-Monats-Wachstumsraten (definiert als $x_t = (z_t - z_{t-3}) / z_{t-3}$, z_t Preis in t) in den Zeitpunkten $t - 1$ und $t - 9$ dargestellt. Der starke Saisoneffekt spiegelt sich im hohen „t-Wert" für den Einfluß aus $t - 9$ wider. (Die damals beobachtete Wachstumsrate korrespondiert zu dem Zeitraum, für den die Preiserwartung in t abgegeben wird.) Der hohe Wert für das χ^2_{PE}-Maß zeigt allerdings, daß die Schätzung insgesamt gesehen wenig vertrauenswürdig ist.

Tabelle 2.2 bezieht sich auf die Produktionspläne in Abhängigkeit von den Vormonatswerten und dem Geschäftsklima; letzteres ist ein Konglomerat aus einer Beurteilung der gegenwärtigen und der zukünftigen Geschäftssituation in *dieser* Produktgruppe (*Strigel* 1972). Das bereits in Abschnitt 1.2 erwähnte Problem der Berücksichtigung von „endogenen verzögerten" Variablen als Einflußgrößen bei aggregierten Tendenzdaten wird hier deutlich: Würde man die Daten für alle drei Kategorien in der Schätzung benutzen, so würde die Summe bis auf einen Proportionalitätsfaktor gleich dem Absolutglied sein, d. h. es bestände exakte Multikollinearität. Ich habe die „−"-Kategorie fortgelassen, doch wäre wohl eine symmetrische Behandlung (Zuordnung der Werte -1, 0, $+1$ zu den drei Kategorien) sinnvoller. Die Tabelle 2.2 zeigt den auch für Mikrodaten gefundenen starken Zusammenhang der Produktionspläne mit den Geschäftserwartungen[13] bzw. mit dem daraus gewonnenen Geschäftsklima. Aber auch der Einfluß der Vormonatswerte für die Produktionspläne ist deutlich sichtbar. Die Güte der Anpassung ist hier besser als in Tabelle 2.1, jedoch immer noch unbefriedigend, zumindest im Sinne der Testtheorie unter der Annahme der in Abschnitt 2.2d genannten Verteilungseigenschaft.

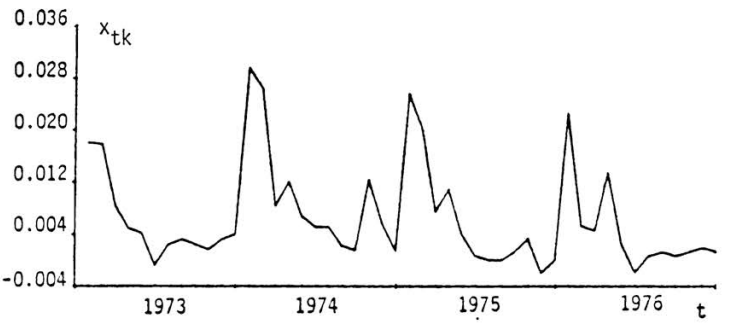
2.4 Simulationsergebnisse

Um einen Anhalt dafür zu bekommen, inwieweit Stichprobenvariation und inwieweit eine Fehlspezifikation bei Annahme des Modells (2-1) für die unbefriedigenden empirischen Ergebnisse verantwortlich

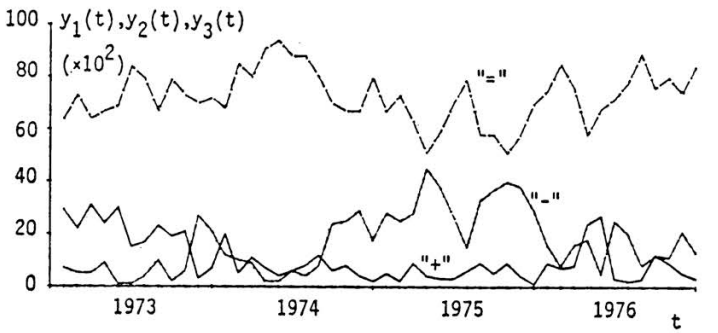
[13] Siehe *König* und *Nerlove* (1980, 209).



a) Preiserwartungen (aggregierte Tendenzdaten)



b) 1-Monats-Wachstumsraten der Verkaufspreise



c) Produktionspläne (aggregierte Tendenzdaten)

Abb. 2/2: Zeitreihen für die Produktgruppe „Werkzeugmaschinen spanabhebend"

ist, wurden anhand von drei Beispielen Simulationsexperimente durchgeführt. (Dabei hat mich, wie eingangs bereits erwähnt, Herr *Arnold Holz* dankenswerterweise tatkräftig unterstützt.) Zwei Beispiele benutzen die exogene Datenkonstellation aus den Tabellen 2.1 und 2.2, im zweiten Fall jedoch nur das Geschäftsklima, nicht die verzögerten Produktionspläne, und geben die (gerundeten) Schätzwerte als wahre Parameter vor. Damit bekommen wir einen Eindruck davon, wie gut die geschätzten Werte das wahre Modell charakterisieren, wenn keine Fehlspezifikation vorliegt, also (2-1) und das angenommene Stichprobenmodell gelten. Das Beispiel I ist fingiert, es ist besonders einfach und dient als Vergleichsmaßstab.

*Tabelle 2.1**

Schätzung des logistischen Modells (2-1) zur Erklärung von Preiserwartungen (aggregierte Tendenzdaten, IFO-KT)

Antwortkategorie	Schätzmethode	geschätzte Koeffizienten der Einflußvariablen		
		Absolutglied	3-monatige Wachstumsraten der Verkaufspreise	
			$t-1$	$t-9$
		$\tilde{\beta}_{1j}$	$\tilde{\beta}_{2j}$	$\tilde{\beta}_{3j}$
$j=3$ („+“)	*ML*	0,26 (1,44)	1,24 (0,22)	27,15 (5,11)
	*BTM*1	0,17 (0,67)	0,21 (0,02)	21,09 (2,96)
$j=2$ („=“)	*ML*	2,26 (13,17)	6,29 (1,14)	− 18,71 (− 3,57)
	*BTM*1	1,81 (7,54)	4,67 (0,64)	− 16,29 (− 2,32)

$\chi^2_{ML} = 399$, $\chi^2_{BTM} = 354$, kritischer Wert: $\chi^2_{90}(0{,}05) = 113$

*) Bemerkungen zu den Tabellen 2.1 und 2.2: *Quelle: Ronning* (1980, Tabellen 3.2 und 3.4). Tabelle 3.2 ist in der endgültigen Fassung nicht mehr erhalten. Die Werte in Klammern geben „*t*-Werte“ an. Für beide Schätzmethoden (*ML* und *BTM*1 (2-16a)) werden die Ergebnisse für die symmetrische Normierung (2-5) angegeben. Für die *BTM*-Schätzung wird dabei (2-8) und (2-9) benutzt. In beiden Tabellen gilt $n(t) = 63$, $t = 1, \ldots, T$. Die χ^2-Maße sind durch (2-24) definiert (Pearson-Maß). Ergebnisse für die *ML*-Schätzung sind gegenüber der genannten Quelle teilweise revidiert (Programmierfehler).

*Tabelle 2.2**

Schätzung des logistischen Modells (2-1) zur Erklärung von Produktionsplänen (aggregierte Tendenzdaten, IFO-KT)

Antwortkategorie	Schätzmethode	geschätzte Koeffizienten der Einflußvariablen: Absolutglied	Produktionspläne im Vormonat („+")	Produktionspläne im Vormonat („=")	Geschäftsklima
		$\tilde{\beta}_{1j}$	$\tilde{\beta}_{2j}$	$\tilde{\beta}_{3j}$	$\tilde{\beta}_{4j}$
$\frac{j=3}{(„+")}$	ML	− 0,9427 (− 2,11)	0,0234 (3,64)	0,0084 (1,72)	0,0228 (7,30)
	$BTM1$	− 0,7833 (− 1,50)	0,0191 (2,40)	0,0068 (1,18)	0,0198 (5,60)
$\frac{j=2}{(„=")}$	ML	0,8403 (2,87)	− 0,0061 (− 1,40)	0,0095 (2,93)	0,0062 (2,84)
	$BTM1$	0,7244 (2,17)	− 0,0091 (− 1,77)	0,0098 (2,66)	0,0061 (2,53)

$\chi^2_{ML} = 181$, $\chi^2_{BTM} = 174$, kritischer Wert: $\chi^2_{88}(0{,}05) = 111$

* Siehe Bemerkungen auf S. 195.

Bei den Simulationsexperimenten ging ich wie folgt vor: Für $t = 1, \ldots, T$ wurden für Wahrscheinlichkeiten (2-1) mit vorgegebenen Werten für den exogenen Vektor $x(t)$ und Parametervektor β jeweils $n(t)$ multinomialverteilte Zufallszahlen erzeugt.

Für alle T Zeitpunkte insgesamt wurden die Zufallszahlen als Beobachtungswerte einer qualitativen abhängigen Variablen aus einer Mehrfach-Multinomial-Stichprobe interpretiert. Die Koeffizienten wurden dann mittels der in Abschnitt 2.2 beschriebenen Schätzmethoden (ML und BTM) ermittelt. Das Verfahren wurde 20mal durchgeführt (= ein Simulationsexperiment). Für jedes der drei Beispiele wurden Simulationsexperimente mit unterschiedlich großem $n(t)$ durchgeführt. Eine kurze Charakterisierung der drei Beispiele ist in Tabelle 2.3 gegeben.

Tabelle 2.3

Charakterisierung der drei Beispiele für die Simulationsexperimente

Experiment No.	Anzahl Kategorien (r)	Anzahl der exogenen Variablen (K)	Anzahl Zeitpunkte (T)	Beobachtungen pro Zeitpunkt ($n(t)$)	Bemerkungen
I	2	2	4	5 10 30 60 100	fingiertes Beispiel; Regressorwerte: 1,0; 1,2; 1,5; 2,0. $\beta_{11} = 4{,}83$; $\beta_{21} = -3{,}30$
II	3	2	48	5 25 60	Geschäftsklima als Einflußvariable. Schätzwerte aus Tabelle 2.2 Siehe auch Abb. 2/1.
III	3	3	48	5 25 60	Einflußvariable aus Tabelle 2.1 mit gerundeten Schätzwerten als Parametern

Für jedes $n(t)$ und jedes Beispiel wurden 20 Läufe durchgeführt.

Da in der *ML*-Methode die Möglichkeit der Nichtkonvergenz besteht, sei im folgenden die Anzahl der „erfolgreichen" Läufe mit G bezeichnet. Für *BTM* gilt stets $G = 20$. Ferner sei β^*_{kj} der wahre Parameter und $\widetilde{\beta}^*_{kj}$ der geschätze Wert (*ML* bzw. *BTM*). Ferner bezeichne q die Werte aus dem q-ten Lauf, $q = 1, \ldots, G$, wobei im folgenden die Symbolik, die bisher gebräuchlich war, um diesen Index erweitert wird. Für jedes Simulationsexperiment wurden folgende Werte bestimmt (Summierung jeweils für $q = 1, \ldots, G$):

„Schätzwert"	$\frac{1}{G} \sum \widetilde{\beta}_{kj}(q) =: \bar{\widetilde{\beta}}_{kj}$
„Standardabweichung"	$\frac{1}{G} \sum s_{kj}(q), s^2_{kj}$
	Diagonalelement in der asymptotischen Kovarianzmatrix

„t-Wert"	$\frac{1}{G} \Sigma \, \tilde{\beta}_{kj}(q) / s_{kj}(q)$
„RMSE-Wert" (root mean-squared error)	$\left[\frac{1}{G} \Sigma \, (\beta^*_{kj} - \tilde{\beta}_{kj}(q))^2\right]^{1/2}$
„Bias"	$\frac{1}{G} \Sigma \, \lvert (\beta^*_{kj} - \tilde{\beta}_{kj}(q)) \rvert$
„relative Häufigkeit"	$\frac{1}{G} \Sigma \, m_j(t, q) / n(t)$
„geschätzte Wahrscheinlichkeit"	$\frac{1}{G} \Sigma \, \tilde{p}_j(t, q)$ mit $\tilde{p}_j(t)$ aus (2-1) und $\bar{\tilde{\beta}}_{kj}$ anstelle von β_{kj}
„χ^2-Testwert"	$\frac{1}{G} \Sigma \, \chi^2_{PE}(q) =: \overline{\chi^2_{PE}}$ bzw. $\frac{1}{G} \Sigma \, \chi^2_{LR}(q) =: \overline{\chi^2_{LR}}$
„Kolmogoroff-Testwert"	Auf der Basis der G verschiedenen χ^2-Werte wird das Kolmogoroff-Prüfmaß unter H_0:
D_G	„χ^2_{PE} bzw. χ^2_{LR} ist χ^2-verteilt mit $(T - K)$ (r − 1) Freiheitsgraden" berechnet.

Da die Simulations-Ergebnisse insgesamt[14] sehr umfangreich sind, soll hier nur über die wichtigsten Aspekte berichtet werden. Diese betreffen einmal die Güte der Anpassung und zum anderen die Schätzung der einzelnen Parameter.

a) Verteilungseigenschaft der χ^2-Maße und Güte der Anpassung

Die Ergebnisse bezüglich der Güte der Anpassung sind für alle drei Beispiele in Tabelle 2.4 zusammengefaßt. Wenn die *Verteilungseigenschaft* für die beiden Anpassungsmaße χ^2_{PE} und χ^2_{LR} (siehe 2.2d) zutreffen soll, dann müssen die arithmetischen Mittel jeweils annähernd gleich der Anzahl der Freiheitsgrade (*FG*) sein. Dies ist nur im Beispiel I für alle Stichprobenumfänge (annähernd) gegeben. Eine exakte Aussage liefert erst die Betrachtung des Kolmogoroff-Prüfmaßes $\tilde{D}_G$: Für Beispiel I sowie für Beispiel III bei großem Stichprobenumfang wird die Verteilungshypothese für das Pearsonmaß χ^2_{PE} bestätigt. Dagegen ist für das Likelihoodquotientenmaß χ^2_{LR} auch bei größeren Stichprobenumfängen die Verteilungseigenschaft weniger deutlich gegeben. Dies ist angesichts der vorherrschenden Präferenz für das letztere Maß (*Bock* 1975, 527) ein wichtiges Ergebnis. Auf der anderen Seite muß darauf hingewiesen werden, daß in Tabelle 2.4 die in Abschnitt 2.2d genannte „Faustregel" $n(t)\,\tilde{p}_j(t) \geq 5$ unbeachtet blieb. Anhand der

[14] Sie können beim Verfasser angefordert werden.

Tabelle 2.4

Zusammenfassung einiger Simulationsergebnisse aus den drei Beispielen

Exp. No.	$n\,(t)$	FG[a]	Schätz-methode	$\overline{\chi^2_{PE}}$	$\tilde{D}_G$[b]	$\overline{\chi^2_{LR}}$	$\tilde{D}_G$[b]	G[c] ($G=20$ für BTM)	Rechen-zeit (TR440-sec)
I	5	2	*ML*	1,85	0,206***	2,24	0,174***	12	19,18
			BTM1	2,33	0,410 *	3,93	0,622		13,38
			BTM2	2,23	0,417 *	3,74	0,628		13,80
	10		*ML*	1,40	0,217***	1,70	0,157***	19	16,07
			BTM1	1,92	0,397	3,13	0,592		13,28
			BTM2	1,92	0,388	3,13	0,589		13,68
	30		*ML*	1,83	0,147***	2,23	0,214***	20	15,10
			BTM1	1,89	0,215***	2,87	0,282 *		13,26
			BTM2	1,90	0,219***	2,88	0,282 *		13,47
	60		*ML*	1,98	0,242***	2,18	0,321 *	20	15,05
			BTM1	2,05	0,260 **	2,64	0,388		13,29
			BTM2	2,05	0,259 **	2,64	0,388		13,68
	100		*ML*	2,38	0,258 **	2,70	0,250 **	20	15,14
			BTM1	2,43	0,288 **	3,14	0,322 *		13,64
			BTM2	2,42	0,293 **	3,13	0,319 *		14,11
II	5	92	*ML*	9,09	1,000	7,10	1,000	19	87,78
			BTM1	60,15	0,961	94,05	0,341 *		33,71
			BTM2	23,23	1,000	37,45	1,000		31,86
	25		*ML*	96,20	0,900	12,23	1,000	20	67,56
			BTM1	123,83	0,905	199,78	1,000		33,50
			BTM2	50,21	0,996	81,09	0,440		32,44
	60		*ML*	46,50	0,900	16,01	1,000	20	64,01
			BTM1	182,27	1,000	294,69	1,000		32,72
			BTM2	64,67	0,898	100,54	0,428		32,37
III	5	90	*ML*	69,37	0,626	58,86	0,947	13	211,65
			BTM1	57,95	0,941	84,69	0,271***		38,42
			BTM2	56,67	0,906	79,37	0,399 *		32,46
	25		*ML*	84,02	0,286 **	69,13	0,658	20	114,39
			BTM1	68,10	0,734	94,25	0,292 **		38,52
			BTM2	87,79	0,523	121,41	0,676		39,19
	60		*ML*	92,68	0,184***	89,47	0,111***	20	102,97
			BTM1	80,80	0,385	111,73	0,671		38,01
			BTM2	87,63	0,166***	126,95	0,810		38,33

a) (Theoretische) Anzahl Freiheitsgrade in den beiden Anpassungstests ($FG = (T - K)(r - 1)$).

b) Kolmogoroff-Smirnoff-Prüfmaß für die G verschiedenen Werte von χ^2_{PE} bzw. χ^2_{LR} unter Zugrundelegung von FG Freiheitsgraden. *** (**; *) bedeutet, daß die Nullhypothese für $\alpha = 0{,}20$ (0,05; 0,01) akzeptiert wird.

c) Anzahl erfolgreicher Simulationsläufe. Für beide BTM-Versionen ergab sich stets $G = 20$.

„geschätzten Wahrscheinlichkeiten" (siehe oben) kann man zeigen, daß die Bedingung vor allem bei den Beispielen II und III oftmals verletzt ist. Einziger sinnvoller Ausweg wäre die Verwendung von „exakten" Tests. Siehe dazu *Nerlove* und *Press* (1978, 46/47).

Vergleicht man die *Güte der Anpassung* der drei Schätzmethoden anhand der χ^2-Maße, so ergibt sich für beide Maße in den Beispielen I und II (mit einer Ausnahme) identische Rangfolge für jeweils ein bestimmtes $n(t)$, ohne daß eine der drei Schätzmethoden überall den „besten Fit", d. h. das kleinste χ^2-Maß, besäße. In Beispiel III differieren die Rangfolgen auch zwischen den beiden Maßen. Vor allem bei diesem Vergleich ist die *geringe Anzahl der Simulationsläufe* zu beachten. Immerhin zeigt Tabelle 2.4 folgendes:

α) $\overline{\chi^2_{LR}}$ ist in allen Beispielen für alle Stichprobenumfänge für die *ML*-Schätzung minimal. Dies bestätigt die theoretische Beziehung zwischen *ML*-Schätzung und Likelihoodquotienten-Test im Vergleich mit anderen Schätzmethoden.

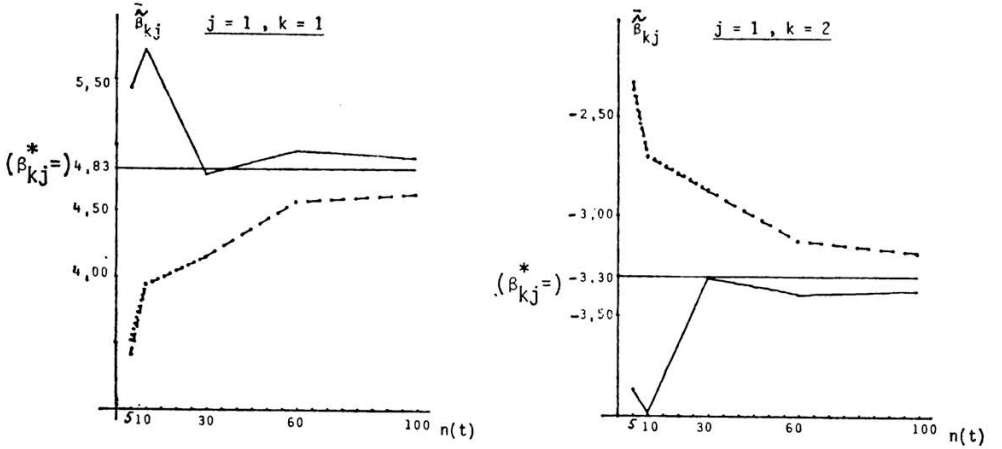
β) In den Beispielen I und II ergibt die *ML*-Schätzung auch für das Pearson-Maß minimale Werte (Ausnahme: Beispiel II, $n(t) = 25$).

γ) Vor allem für kleine Stichprobenumfänge schneidet die *BTM*2-Schätzung gegenüber der in der Literatur üblichen *BTM*1-Schätzung bezüglich des „Fits" besser ab.

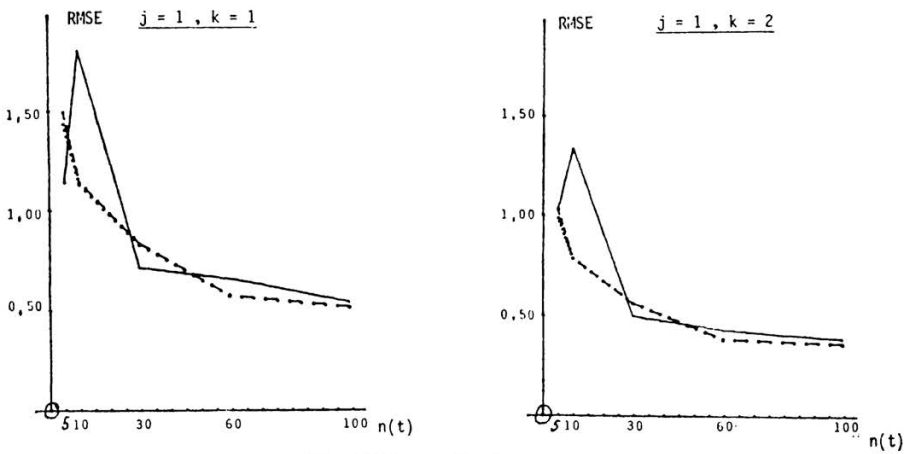
Schließlich sei auf die in der Tabelle ebenfalls aufgeführte Anzahl erfolgreicher Simulationsläufe hingewiesen: Für kleine Stichprobenumfänge ist vor allem in den Beispielen I und III die „Abbruchrate" recht groß. Man beachte, daß die abgebrochenen Läufe bei der Analyse der Simulationsergebnisse unberücksichtigt bleiben. Die Frage, ob sich daraus eine Verzerrung ergibt, kann meines Wissens nicht beantwortet werden.

b) Schätzung der Koeffizienten

Für *Beispiel I* zeigt die Abbildung 2/3 die Simulationsergebnisse bezüglich der geschätzten Koeffizienten und die daraus resultierenden *RMSE*-Werte. Weil wir es hier nur mit $r = 2$ Kategorien zu tun haben, können wir uns wegen der Normierung (siehe Abschnitt 2.2a) auf die Ergebnisse für *eine* Kategorie beschränken. Die Abbildung zeigt, daß die beiden *BTM*-Versionen in diesem Beispiel nahezu identische Schätzwerte (und folglich auch *RMSE*-Werte) ergeben. Für wachsenden Stichprobenumfang $n(t)$ nähern sich für alle drei Methoden die Schätzwerte den jeweiligen wahren Werten an. Auffallend ist die starke Abweichung der *ML*-Schätzung für $n(t) = 10$ in Abbildung 2/3a.



a) Schätzwerte (wahre Werte: $\beta_{11} = 4{,}83$, $\beta_{21} = -3{,}30$)



b) RMSE - Werte

Abb. 2/3: Simulationsergebnisse für Beispiel I
(——— *ML*, - - - - *BTM*1, ... *BTM*2)

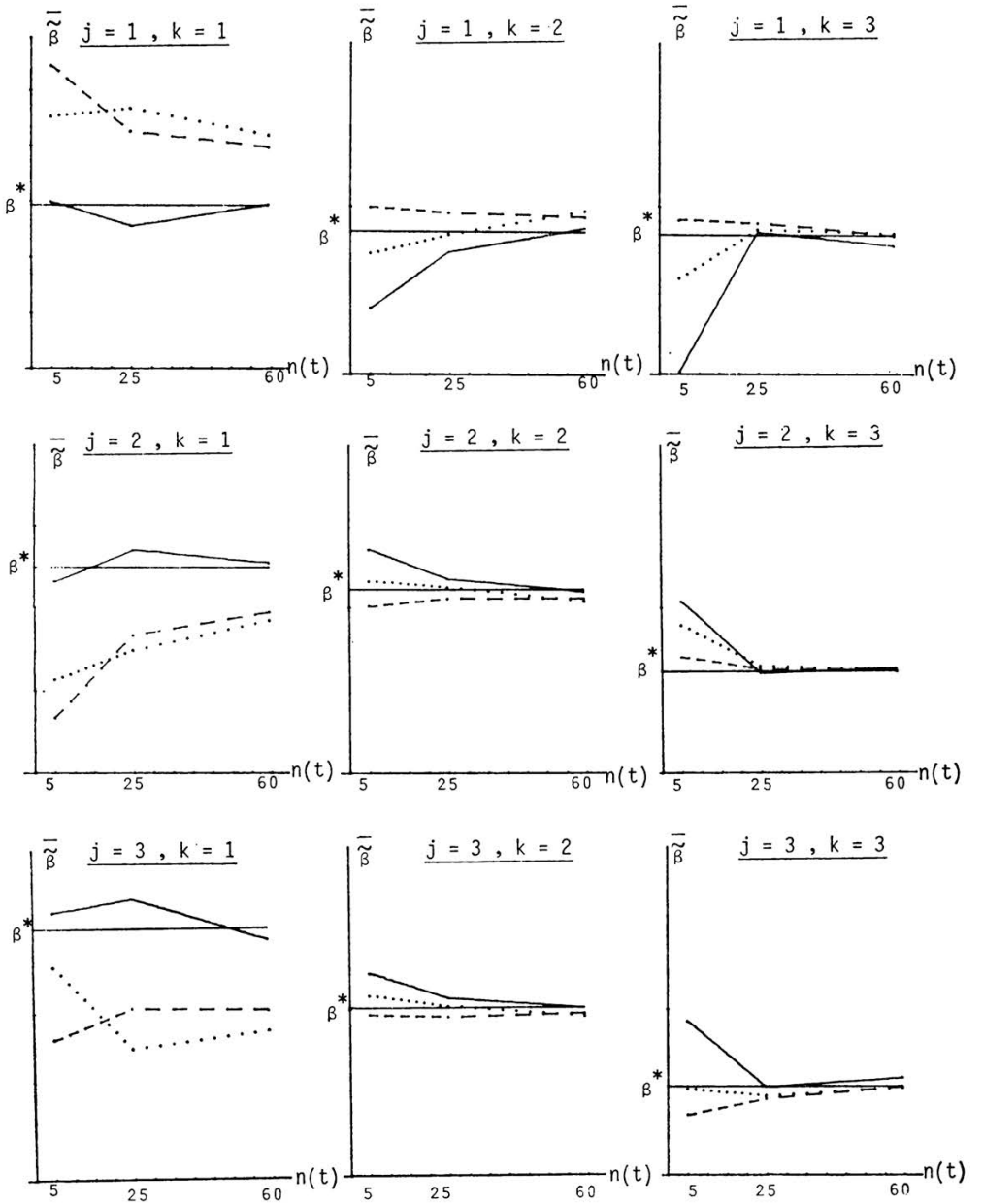


Abb. 2/4: Simulationsergebnisse für Beispiel III.
Schätzwerte (——— *ML*, - - - - *BTM*1, ... *BTM*2)

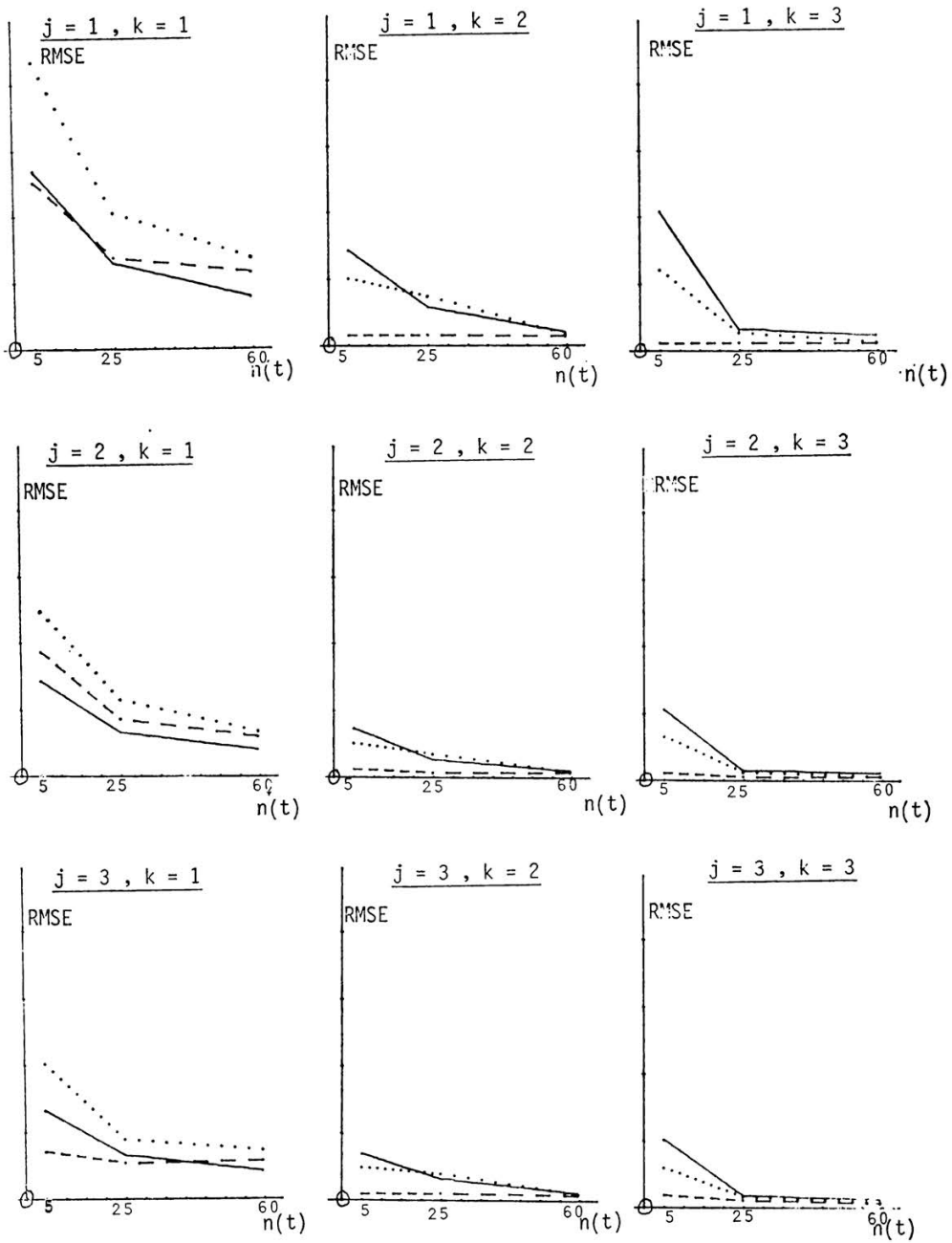


Abb. 2/5: Simulationsergebnisse für Beispiel III.
RMSE-Werte (——— *ML*, - - - - *BTM*1, ... *BTM*2)

Für Beispiel III[15] sind die Ergebnisse aus den Simulationen in den Abbildungen 2/4 und 2/5 für alle drei Kategorien graphisch dargestellt, wobei die durch die Normierung bedingten Restriktionen zu beachten sind. Abbildung 2/2 zeigt, daß bei der Schätzung des „Absolutgliedes" ($k = 1$) die *ML*-Schätzung am besten abschneidet, während sich — vor allem für kleinere Stichprobenumfänge — bei den „Strukturkoeffizienten" ($k = 2, 3$) eine deutliche Überlegenheit der *BTM*-Schätzung ergibt. Bezüglich der *RMSE-Werte* (siehe Abbildung 2/5) schneidet bei allen drei Einflußvariablen *BTM*1 deutlich besser ab als *BTM*2.

c) Restriktionen für geordnete Kategorien

Wie bereits in Abschnitt 2.1 besprochen wurde, hat in logistischen Modellen, in denen nur *eine* Einflußvariable ($K = 2$) spezifiziert ist, die *Rangordnung* der Schätzwerte für die Koeffizienten dieser Variablen in den einzelnen Kategorien Informationswert. Dies soll hier kurz illustriert werden: Für Beispiel II, $n(t) = 5$, ergaben sich folgende „Schätzwerte" in den Simulationen:

Tabelle 2.5

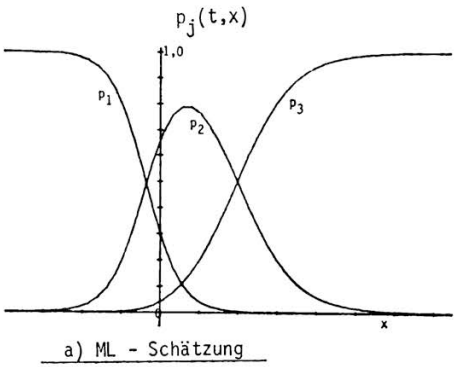
Schätzwerte für β_{2j} im Beispiel II, n (t) = 5 (auf der Basis von simulierten Werten)

Schätzmethode	$j = 1$	$j = 2$	$j = 3$
ML	— 3,90	0,56	3,34
*BTM*1	— 0,28	0,23	0,05
*BTM*2	— 0,76	0,31	0,45
„wahre Werte" ..	— 2,90	0,60	2,30

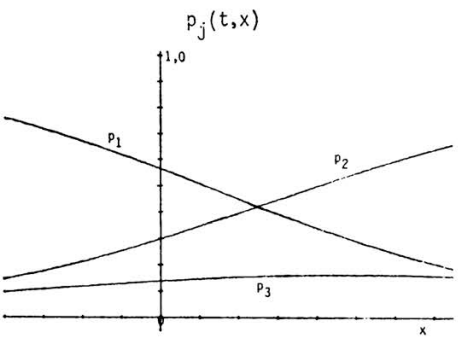
Gemäß (2-3) sollten die Koeffizienten für die drei Kategorien der Größe nach geordnet sein[16]. Dies ist für die „wahren Werte" sowie für *ML* und *BTM*2 der Fall, nicht dagegen für *BTM*1. Die Schätzwerte dieser Methode können also bereits aufgrund der falschen Rangordnung verworfen werden. Darüber hinaus zeigt Abbildung 2/6, in der die aus

[15] Die Simulationsergebnisse für Beispiel II (Schätzwerte und RMSE-Werte) können hier aus Platzgründen nicht gezeigt werden. In diesem Beispiel schneiden die *ML*-Schätzwerte am besten ab. (Sie sind auch, wie unten im Zusammenhang mit Tabelle 2.5 erläutert wird, selbst für kleine Stichprobenumfänge sinnvoll.) Allerdings ist die Überlegenheit der *ML*-Methode bezüglich der *RMSE*-Werte weniger deutlich.

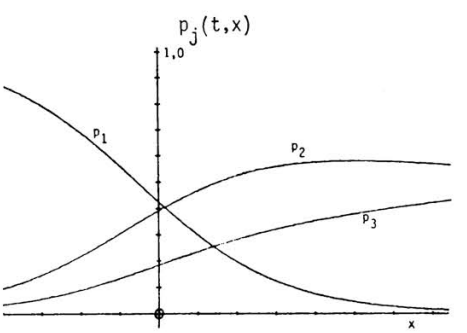
[16] Dabei wird unterstellt, daß die in Abbildung 2/1 gezeigten typischen



a) ML - Schätzung



b) BTM1 - Schätzung



c) BTM2 - Schätzung

Abb. 2/6: Geschätzte Antwortwahrscheinlichkeiten für Beispiel II, $n(t) = 5$ (auf der Basis von simulierten Werten).

den Schätzwerten resultierenden Antwortwahrscheinlichkeiten graphisch dargestellt sind, daß nur die *ML*-Methode aussagekräftige Größenordnungen der Koeffizienten aufweist. Man vergleiche dazu auch Abbildung 2/1, in der die Antwortwahrscheinlichkeiten für die „wahren Werte" aus Tabelle 2.5 dargestellt sind.

3. Schlußbemerkungen

Das logistische Modell (2-1) ist grundsätzlich ein sinnvoller Ansatz zur Erklärung von aggregierten Tendenzdaten. Allerdings zeigen die bisherigen empirischen Ergebnisse (*Ronning* 1980), die beispielhaft in Abschnitt 2.3 dargestellt sind, eine unbefriedigende Erklärungsgüte. Deshalb wurde in einer Simulationsstudie, über die in Abschnitt 2.4 berichtet wird, untersucht, inwieweit die wenig zufriedenstellenden empirischen Ergebnisse auf eine Fehlspezifikation zurückzuführen sind. Gleichzeitig wurde das Verhalten alternativer Schätzmethoden (*ML*, *BTM*) und Gütemaße (χ^2_{PE} , χ^2_{LR}) betrachtet.

Ein Vergleich der empirischen Ergebnisse mit den Simulationsergebnissen in Tabelle 2.4 zeigt, daß bei Erfüllung der Annahme der „Mehrfach-Multinomial-Stichprobe" als datenerzeugendem Prozeß die Erklärungsgüte — bei identischen Einflußgrößen — deutlich ansteigt. Im Zusammenhang mit Paneldaten dürfte dies — in der Literatur übliche — Stichprobenmodell jedoch oftmals verletzt sein (siehe Abschnitt 2.2). Insofern geben die Simulationsergebnisse keinen Aufschluß darüber, inwieweit eine Fehlspezifikation der verwendeten Einflußgrößen *oder* des unterstellten Stichprobenmodells vorliegt. Eine weitere Untersuchung in dieser Richtung ist geplant.

Die Analyse der Verteilungseigenschaften der beiden χ^2-Gütemaße in der Simulationsstudie deutet allerdings an, daß für die betrachteten Stichprobenumfänge (und Parameterkonstellationen), die denen der empirischen Analyse entsprechen, die in der Literatur empfohlene Testprozedur häufig unangemessen ist: Beide Maße erfüllen auf der Basis des Kolmogoroff-Anpassungs-Tests nur in Ausnahmefällen die χ^2-Verteilung mit $(T - K)(r - 1)$ Freiheitsgraden. Siehe dazu Abschnitt 2.2d. Es zeigt sich auch, daß im Fall der *BTM*-Schätzung das Likelihoodquotienten χ^2_{PE} weniger gut als das Pearson-Maß χ^2_{LR} geeignet ist, während im Fall der *ML*-Schätzung das erstgenannte zumindest bei großen Stichproben vorzuziehen ist.

Verläufe der Antwortwahrscheinlichkeiten im Zusammenhang mit der hier betrachteten Einflußvariablen (Geschäftsklima) ebenfalls sinnvoll interpretierbar sind, was sicher der Fall ist.

Der Vergleich der genannten Schätzmethoden in der Simulationsstudie ergibt, daß auch für den bisher weniger betrachteten polytomen (hier: trichotomen) Fall die *BTM*-Schätzung neben der *ML*-Schätzung bestehen kann. Dies gilt vor allem, wie ausgeführt, bei Benutzung von aggregierten Daten auf der Basis von „mittleren" Stichprobenumfängen. Dabei wird neben der in der Literatur üblicherweise verwendeten Version der Berkson-Theil-Methode („*BTM*1") gleichzeitig eine andere Version („*BTM*2") benutzt. Siehe dazu (2-16). Die Simulationsergebnisse zeigen, daß dies eine diskussionswürdige Alternative ist, die allerdings weiterer Untersuchungen vor einer abschließenden Bewertung bedarf.

Die Analyse dieser Arbeit beschränkt sich auf den *univariaten* Fall, bei IFO-Tendenzdaten also auf *eine* Frage aus dem Konjunkturtest. Für den Fall von Mikrodaten bestehen bereits multivariate logistische Modelle zur Erklärung von IFO-Tendenzdaten für mehrere Fragen gemeinsam (*König* und *Nerlove* 1980). Die Frage der Modellierung im Fall *aggregierter* Tendenzdaten stellt sich auch hier.

Zusammenfassung

Das logistische Modell ist ein sinnvoller Ansatz zur Erklärung von aggregierten Tendenzdaten durch quantitative und qualitative Einflußgrößen. Da empirische Resultate auf der Basis von aggregierten IFO-Tendenzdaten für Preiserwartungen und Produktionspläne (*Ronning* 1980) unbefriedigende Erklärungsgüte aufwiesen, wird in dieser Arbeit mittels einer Simulationsstudie die Frage behandelt, ob die Spezifikation der (falschen) Einflußgrößen oder das zugrundegelegte stochastische Modell für den datenerzeugenden Prozeß diese wenig zufriedenstellenden Ergebnisse bedingen. Die Simutionen offenbaren, daß die Annahme der „Mehrfach-Multinominal-Stichprobe", die in den empirischen Untersuchungen unterstellt wurde, dort unangemessen ist, weil sie nicht den Zeitreihen-Charakter der aggregierten Tendenzdaten berücksichtigt.

Es werden zwei alternative Schätzmethoden (Maximum-Likelihood-Methode und Berkson/Theil-Methode) sowie zwei Gütemaße (Pearson-Chiquadrat und Likelihood-Verhältnis) betrachtet. Die Ergebnisse der Simulationsstudie deuten darauf hin, daß für aggregierte Tendenzdaten, die auf „mittleren" Stichprobenumfängen basieren, die Berkson/Theil-Methode und das Pearson-Maß verwendet werden sollten.

Summary

It is shown that the explanation of aggregated tendency data by other quantitative or qualitative variables can be reasonably performed by use of the logistic model. Since empirical results based on aggregated IFO tendency data concerning price expectations and production plans (*Ronning* 1980) showed only small explanatory power, the simulation study presented in this paper was designed to answer the question whether the unsatisfactory

results were due to the (wrong) specification of explanatory variables or to the stochastic model assumed as the data generating process. The results of the simulation study reveal that the assumption of "product multinominal sampling" which was used in empirical research, causes the main trouble, since the assumption does not recognize the time series character of the aggregated tendency data.

Two alternative estimation methods (Maximum Likelihood and the approach proposed by Berkson and Theil) as well as two measures of goodness of fit (Pearson Chi-Square and Likelihood Ratio) were considered. The simulation study indicates that for aggregated tendency data based on a moderate sample size, the Berkson/Theil approach and the Pearson measure should be chosen.

Literatur

Amemiya, T. (1975), Qualitative Response Models, Annals of Economic and Social Measurement 4, 363 - 372.

— (1976), The Maximum Likelihood, the Minimum Chi-Square and the Nonlinear Weighted Least-Squares Estimator in the General Qualitative Response Model, Journal of the American Statistical Association 71, 347 - 351.

Bishop, Y. M. M., S. E. *Fienberg* und P. W. *Holland* (1975), Discrete Multivariate Analysis. Theory and Practice, Cambridge (Mass.).

Bock, R. D. (1975), Multivariate Statistical Methods in Behavioral Research, New York.

Dhrymes, P. J. (1978), Introductory Econometrics, New York.

Domencich, T. A. und D. *McFadden* (1975), Urban Travel Demand, Amsterdam.

Haberman, S. J. (1974), The Analysis of Frequency Data, Chicago.

— (1978), Analysis of Qualitative Data. Vol. I: Introductory Topics, New York.

Heckman, J. J. (1978), Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence against the Hypothesis of Spurious State Dependence, Annales de l'INSEE 30/31, 227 - 269.

INSEE (1978), The Econometrics of Panel Data, Annales de l'INSEE 30/31.

König, H. und M. *Nerlove* (1980), Micro-Analysis of Realizations, Plans and Expectations in the IFO Business Test by Multivariate Log-Linear Probability Models, in: W. Strigel (Hrsg.): Business Cycle Analysis. Proceedings of the 14th CIRET Conference, Westmead, 187 - 226.

— *M. Nerlove* und *G. Oudiz* (1979), Modèles log-linéaires pour l'analyse des données qualitatives: Application a l'étude des enquêtes de conjuncture de l'INSEE et de l'IFO, Annales de l'INSEE 36, 31 - 83.

Lee, T. C., G. G. *Judge* und *A. Zeller* (1977), Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data, 2. A., Amsterdam.

Nerlove, M. und S. J. *Press* (1973), Univariate and Multivariate Log-Linear and Logistic Models, Rand Monograph R-1306-EDA/NIH.

— — (1978), Multivariate Log-Linear Probability Models for the Analysis of Qualitative Data, Manuskript (3. Druck), Northwestern University, Center for Statistics and Probability.

Ronning, G. (1980), Logit, Tobit and Markov Chains: Three Different Approaches to the Analysis of Aggregated Tendency Survey Data, in: W. Strigel (Hrsg.): Business Cycle Analysis. Proceedings of the 14th CIRET-Conference, Westmead, 227 - 257.

Schneeweiß, H. (1978), Ökonometrie, 3. A., Würzburg.

Strigel, W. (1972), Konjunkturindikatoren aus qualitativen Daten, IFO-Studien 18, 185 - 214.

Theil, H. (1967), Economics and Information Theory, Amsterdam.

— (1970), On the Estimation of Relationships Involving Qualitative Variables, American Journal of Sociology 76, 103 - 154.

Vogler, K. (1978), Content and Determinants of Judgemental and Expectational Variables in the IFO Business Survey, in: W. Strigel (Hrsg.): Problems and Instruments of Business Cycle Analysis, Berlin, 75 - 114.