

# Alternative Strukturen des Nested Logit Modells mit der PROC MDC<sup>1</sup>

Nadja Silberhorn  
Humboldt-Universität zu Berlin  
Institut für Marketing  
Spandauer Straße 1  
10178 Berlin  
silberhorn@wiwi.hu-berlin.de

Till Dannewald  
Humboldt-Universität zu Berlin  
Institut für Marketing  
Spandauer Straße 1  
10178 Berlin  
dannewald@wiwi.hu-berlin.de

## Zusammenfassung

Der Beitrag stellt verschiedene Spezifikationen des Nested Logit Modells vor und demonstriert deren Umsetzung mit Hilfe der PROC MDC.

Das Nested Logit Modell wird wegen seiner Fähigkeit, Ähnlichkeiten zwischen Alternativen zuzulassen und abzubilden, zunehmend in praktischen Anwendungen vor allem in den Bereichen Logistik und Marketing herangezogen. Allerdings wird der Tatsache, dass zwei unterschiedliche Formen des Nested Logit Modells existieren, kaum Bedeutung geschenkt. Aus vielen Veröffentlichungen geht nicht hervor, mit welcher Spezifikation gerechnet wurde. Die Vergleichbarkeit der Ergebnisse ist somit nicht zwingend gewährleistet. Das UMNL (*utility maximization nested logit*) Modell und das NNNL (*non-normalized nested logit*) Modell haben unterschiedliche Eigenschaften, die sich auf die Schätzergebnisse (Schätzer der Koeffizienten der Nutzenfunktionen, Baumstrukturen), Interpretationen und Prognosen auswirken. Diese Unterschiede können substantiell auf Marktforschungserkenntnisse und Marketingentscheidungen einwirken.

Da die NNNL Struktur nicht konsistent mit der Theorie der Nutzenmaximierung (Koppelman & Wen, 1998a) ist, wird die UMNL Struktur bevorzugt. Der PROC MDC in SAS© 9.1.3 ist nur die NNNL Spezifikation hinterlegt, so dass zur Schätzung entsprechend der Nutzenmaximierung die Einführung von Restriktionen nötig ist.

**Schlüsselworte:** Nested Logit Modell, utility maximization nested logit (UMNL), non-normalized nested logit (NNNL), PROC MDC.

## 1 Einleitung

Zur Modellierung von diskreten Konsumentenentscheidungen wird im Kontext der Nutzenmaximierung standardmäßig das Logit Modell herangezogen. Jedoch geht das Standard Logit Modell von der Annahme proportionaler Substitutionsbeziehungen (IIA) aus, die besagt, dass das Verhältnis der Auswahlwahrscheinlichkeiten zweier Alternativen nicht von der An- oder Abwesenheit anderer Alternativen im Modell abhängig ist. Anhand eines einfachen Beispiels lässt sich dies verdeutlichen. Ausgehend von der Annahme, dass sich die beiden Zeitschriften B und G den Markt für Publikumszeitschriften mit einem Marktanteil von jeweils 50 % aufteilen, verändert

---

<sup>1</sup> Finanziell gefördert durch die Deutsche Forschungsgemeinschaft (DFG) im Forschungsprojekt #BO1952/1 und im Sonderforschungsbereich 649 "Ökonomisches Risiko".

sich bei Gültigkeit der IIA-Annahme das Verhältnis der Marktanteile von B und G durch den Markteintritt der Zeitschrift M nicht. Die Zeitschrift M zieht also gleichermaßen Marktanteile von B und G ab. Es könnten sich für B, G und M Marktanteile von 40 %, 40 % und 20 % ergeben. Dies ist in der Realität so nicht der Fall. Es ist vielmehr davon auszugehen, dass sich das 1:1 Verhältnis der Auswahlwahrscheinlichkeiten respektive der Marktanteile der Zeitschriften B und G verschiebt.

Zur Überwindung der restriktiven IIA-Annahme wurde für die vorliegende Studie der Nested Logit (NL) Ansatz gewählt. Das Nested Logit Modell erlaubt allgemeinere Substitutionsmuster als das Standard Logit Modell und bleibt trotzdem analytisch greifbar. Nested Logit Modelle tragen der Tatsache Rechnung, dass jede Alternative spezifische Informationen in ihrer unbeobachtbaren Nutzenkomponente haben kann, die eine Rolle im Entscheidungsprozess spielen. Die Fehlerkomponenten der Alternativen müssen nicht die gleiche Verteilung haben. Untermengen der Alternativen haben einen ähnlichen Informationsgehalt, so dass Korrelationen zwischen Alternativenpaaren bestehen (Hensher et al., 2005).

Die Einteilung der Alternativen entsprechend ihrer Ähnlichkeiten in Nester und die daraus resultierende Baumstruktur haben nichts gemeinsam mit einer stochastischen Alternativenbewertung im Rahmen eines Entscheidungsbaums. Nested Logit Modelle definieren nicht den Prozess der Entscheidungsfindung, sondern berücksichtigen Varianzunterschiede in den unbeobachtbaren Nutzenkomponenten (Hensher et al., 2005).

Das Nested Logit Modell wird wegen seiner Fähigkeit, Ähnlichkeiten zwischen Alternativen zuzulassen und diese durch die teilweise Korrelation der Fehlerterme abzubilden, zunehmend in praktischen Anwendungen, insbesondere in der Marketingforschung (Chintagunta & Vilcassim, 1998; Ngobo, 2005), herangezogen. Allerdings wird der Tatsache, dass zwei unterschiedliche Formen des Nested Logit Modells existieren, kaum Bedeutung geschenkt. Aus vielen Veröffentlichungen geht nicht hervor, mit welcher Spezifikation gerechnet wurde. Die Vergleichbarkeit der Ergebnisse ist somit nicht gewährleistet. Das *utility maximization nested logit* (UMNL) Modell und das *non-normalized nested logit* (NNNL) Modell haben unterschiedliche Eigenschaften, die sich auf die Schätzergebnisse, Interpretationen und Prognosen auswirken. Diese Unterschiede können substanziell auf Marketingentscheidungen einwirken.

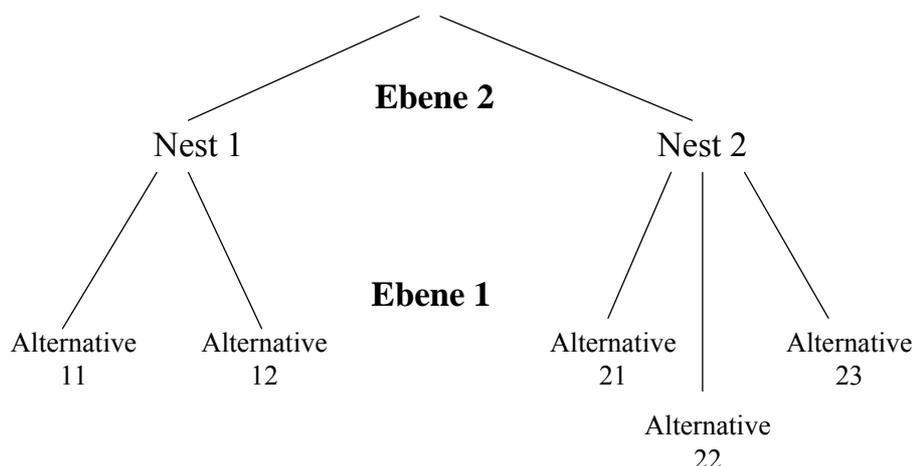
Der *PROC MDC* in SAS© 9.1.3 ist die NNNL Spezifikation, die nicht konsistent mit der Theorie der Nutzenmaximierung ist (Koppelman & Wen, 1998a), hinterlegt, so dass zur Schätzung entsprechend der Nutzenmaximierung die Einführung von Restriktionen nötig ist. Anhand eines fiktiven Beispiels aus dem Verlagswesen werden die Auswirkungen der unterschiedlichen Nested Logit Spezifikationen auf die Schätzergebnisse demonstriert.

Der Beitrag ist wie folgt gegliedert:

Die in Abschnitt 2 vorgestellten alternativen Formen des Nested Logit Modells werden in Abschnitt 3 auf ihre Konsistenz mit der Theorie der Nutzenmaximierung überprüft. In Abschnitt 4 wird anhand eines konstruierten Beispiel aus dem Verlagswesen die Umsetzung der Schätzung eines Nested Logit Modells mit der PROC MDC in SAS© 9.1.3 unter Berücksichtigung der zu Grunde liegenden Spezifikation demonstriert. Der Beitrag schließt in Abschnitt 5 mit einer zusammenfassenden Handlungsempfehlung. Die SAS© 9.1.3 und STATA© 9.1 Programmcodes befinden sich im Anhang.

## 2 Spezifikationen des Nested Logit Modells

Die Entscheidungssituation in einem Nested Logit Modell kann anhand einer Baumstruktur (**Abbildung 1**) verdeutlicht werden. Zur Ermittlung des "besten" Baums im Hinblick auf den Log-Likelihood Wert schlagen Hensher et al. (2005) den "degenerate nested logit" Ansatz vor.



**Abbildung 1:** Baumstruktur eines Nested Logit Modells

Der Zufallsnutzen jeder Alternative  $U_{im}$  ergibt sich aus der Summe einer marginalen Nutzenkomponente  $U_m$  aus Ebene 2 und einer konditionalen Nutzenkomponente  $U_{i|m}$  aus Ebene 1, die jeweils aus einem deterministischen Part  $V$  und einem stochastischen Part  $v$  bestehen.

$$U_{im} = U_m + U_{i|m} = (V_m + v_m) + (V_{i|m} + v_{i|m})$$

Die Fehlerterme  $v_m$  und  $v_{i|m}$  sind unabhängig. Während die Fehlerterme  $v_{i|m}$  identisch und unabhängig (iid) Gumbel-verteilt sind mit dem Skalenparameter  $\mu_m$ , sind die zusammengesetzten Fehlerterme  $\varepsilon_{im} = v_m + v_{i|m}^*$  so verteilt, dass die Summe von  $U_m$  und  $U_{i|m}^*$ , dem Maximum der  $U_{i|m}$ , Gumbel-verteilt ist mit dem Skalenparameter  $\lambda_m$  (Ben-Akiva & Lerman, 1985; Hunt, 2000).

$$Var(v_{i|m}) = \frac{\pi^2}{6\mu_m^2}$$

$$Var(\varepsilon_{im}) = Var(v_m + v_{i|m}^*) = \frac{\pi^2}{6\lambda_m^2}$$

Da die zusammengesetzten unbeobachtbaren Nutzenkomponenten  $\varepsilon_{im}$  Varianzkomponenten der unteren und der oberen Entscheidungsebene beinhalten, können die Varianzen der oberen Ebene nicht kleiner sein als die der unteren. Aus diesem Grund müssen die Skalenparameter der folgenden Bedingung genügen (Carrasco & de Dios Ortúzar, 2002; Hensher et al., 2005):

$$\lambda_m < \mu_m$$

Die erste Spezifikation, das so genannte *non-normalized nested logit* (NNNL) Modell wurde aus dem Standard Logit Ansatz abgeleitet, um die IIA-Annahme zu lockern. Die elementare NNNL-Form ist nicht konsistent mit der Theorie der Nutzenmaximierung. Dem gegenüber steht das *utility-maximizing nested logit* (UMNL) Modell, das aus der "Generalized Extreme Value" (GEV) Theorie von McFadden (McFadden, 1978; McFadden, 1981) abgeleitet wurde und konsistent mit der Theorie der Nutzenmaximierung ist. Die beiden Formen des Nested Logit Modells haben verschiedene Eigenschaften, die die Schätzung beeinflussen und, abhängig von der tatsächlichen Zielfunktion, falsche Ergebnisse hervorbringen können. Marketingentscheidungen, die auf diesen Ergebnissen beruhen, können somit wesentlich beeinträchtigt werden. **Tabelle 1** stellt in Anlehnung an Hunt (2000) die NNNL der UMNL Spezifikation gegenüber.

Die Auswahlwahrscheinlichkeit  $P_{im}$  für Alternative  $im$  ergibt sich aus dem Produkt der marginalen Auswahlwahrscheinlichkeit  $P_m$  für Nest  $m$  und der konditionalen Auswahlwahrscheinlichkeit  $P_{i|m}$  für eine Alternative  $i$  innerhalb des Nests  $m$ . Der "inclusive value"  $IV_m$  ist der erwartete Wert des Nests  $m$  und verbindet die obere mit der unteren Entscheidungsebene.

Der Unterschied zwischen den beiden Nested Logit Formen besteht in der expliziten Skalierung der deterministischen Nutzenkomponente in der UMNL Spezifikation. Im Falle generischer Koeffizienten, d. h. konstanter Koeffizienten über alle Elementaralternativen hinweg, bedeutet dies für die NNNL Spezifikation, dass zwar die geschätzten und ausgewiesenen Parameter für alle Alternativen konstant sind, nicht aber die sich dahinter verbergenden "wahren" Parameter. Dies liegt an der implizit stattgefundenen nestspezifischen Skalierung bei der NNNL Form.

Auf Grund von Identifikationsproblemen ist bei der UMNL Spezifikation einer der beiden Skalenparameter auf den Wert 1 zu normalisieren (Hunt, 2000). Wird eine Normalisierung auf der unteren Ebene ( $\mu_m = \mu_n = 1$ ) gewählt, spricht man vom UMNL RU1 Modell, wird die obere Ebene zur Normalisierung ( $\lambda_m = \lambda_n = 1$ ) gewählt, spricht man vom UMNL RU2 Modell.

**Tabelle 1:** Spezifikationen des Nested Logit Modells

	UMNL	NNNL
$P_{im}$	$P_m \cdot P_{i m}$	$P_m \cdot P_{i m}$
$P_m$	$\frac{\exp\left(\lambda_m V_m + \frac{\lambda_m}{\mu_m} IV_m\right)}{\sum_n \exp\left(\lambda_n V_n + \frac{\lambda_n}{\mu_n} IV_n\right)}$	$\frac{\exp(V_m + \kappa_m IV_m)}{\sum_n \exp(V_n + \kappa_n IV_n)}$
$IV_m$	$\ln \sum_{j \in C_m} \exp(\mu_m V_{j m})$	$\ln \sum_{j \in C_m} \exp(V_{j m})$
$P_{i m}$	$\frac{\exp(\mu_m V_{i m})}{\sum_{j \in C_m} \exp(\mu_m V_{j m})}$	$\frac{\exp(V_{i m})}{\sum_{j \in C_m} \exp(V_{j m})}$

### 3 Konsistenz mit der Theorie der Nutzenmaximierung

Die Zufallsnutzentheorie ist ein wesentlicher Bestandteil des Marketing im Sinne der diskreten Wahlmodelle. Um der Theorie der Nutzenmaximierung zu genügen, darf sich die Auswahlwahrscheinlichkeit einer Alternative durch Hinzufügen eines konstanten Terms  $a$  zur deterministischen Nutzenkomponente einer jeden Alternative nicht verändern (Koppelman & Wen, 1998a). **Tabelle 2** verdeutlicht, dass ohne Auferlegung von Restriktionen nur die RU2 Normalisierung der UMNL Form diesem Anspruch genügt.

Nur bei der RU2 Normalisierung stimmt die Auswahlwahrscheinlichkeit  $P_{im}$  mit der Auswahlwahrscheinlichkeit  $P_{im}^*$  nach Hinzufügen eines Terms  $a$  zur Nutzenkomponente  $V_{i|m}$  überein. Der neue inverse value  $IV_m^*$  entspricht in der NNNL und der UMNL RU1 Spezifikation der Summe aus dem alten inverse value  $IV_m$  und der hinzugefügten Konstante  $a$ . Beim UMNL RU2 Modell ist die hinzugefügte Konstante  $a$  noch zusätzlich mit dem Skalenparameter  $\mu_m$  skaliert. Während sich bei allen drei Nested Logit Spezifikationen die neue Auswahlwahrscheinlichkeit  $P_{i|m}^*$  nicht von der alten Auswahlwahrscheinlichkeit  $P_{i|m}$  unterscheidet, sind bei der neuen Auswahlwahrscheinlichkeit  $P_m^*$  auf dem oberen Level Unterschiede auszumachen. Ohne die Auferlegung von Restriktionen genügt nur die UMNL RU2 Spezifikation dem Anspruch der Konsistenz mit der Nutzentheorie. Denn nur bei der RU2 Form stimmt die Auswahlwahrscheinlichkeit  $P_{im}$  mit der Auswahlwahrscheinlichkeit  $P_{im}^*$  nach Hinzufügen eines Terms  $a$  zur Nutzenkomponente  $V_{i|m}$  überein.

**Tabelle 2:** Nested Logit Formen und Nutzenmaximierung

	UMNL <i>utility maximization nested logit</i>		NNNL <i>non-normalized nested logit</i>
	RU1 $(\mu_m = \mu_n = 1)$	RU2 $(\lambda_m = \lambda_n = 1)$	
$V_{i m}^*$	$V_{i m} + a$	$V_{i m} + a$	$V_{i m} + a$
$IV_m^*$	$IV_m + a$	$IV_m + a\mu_m$	$IV_m + a$
$P_{i m}^*$	$P_{i m}$	$P_{i m}$	$P_{i m}$
$P_m^*$	$\neq P_m$	$P_m$	$\neq P_m$
$P_{im}^*$	$\neq P_{im}$	$P_{im}$	$\neq P_{im}$

Die RU1 Normalisierung der UMNL Form und die NNNL Form können jeweils durch die Einführung von Parameter-Restriktionen in ein mit der Nutzenmaximierungstheorie konsistentes Modell überführt werden. Bei der UMNL RU1 Spezifikation kann die Konsistenz nur durch die Restriktion  $\lambda_m = \lambda_n = \lambda$ , also die Parametergleichheit über alle Nester hinweg, erreicht werden. Die Konsistenz mit der Zufallsnutzentheorie kann bei der NNNL Form nur durch die Restriktion  $\mu_m = \mu_n = \mu$  gewährleistet werden.

Heiss (2002) und Hunt (2000) haben gezeigt, dass die restriktive NNNL Form die Schätzergebnisse der restriktiven, RU1 normalisierten UMNL Form reproduziert. Durch Reskalierung der Parameterschätzer im restriktiven NNNL Modell mit dem geschätzten IV-Parameter erhält man die Parameterschätzer des restriktiven, RU2 normalisierten UMNL Modells. Man darf dabei jedoch nicht vergessen, dass jede Restriktion auf die zu schätzenden Parameter mit einem Informationsverlust einhergeht.

$$NNNL_{res} = RUI_{res}$$

$$NNNL_{res} * IV_{NNLres} = RU2_{res}$$

Koppelman & Wen (1998a) haben eine zweite Möglichkeit aufgezeigt, mit der die Konsistenz des NNNL Modells mit der Nutzenmaximierungstheorie auch ohne IV-Parameter Restriktionen sichergestellt werden kann. Zum einen müssen zusätzliche Dummy-Nester unterhalb der bisherigen unteren Ebene eingeführt werden, zum anderen müssen die dadurch zusätzlich geschätzten Skalenparameter so definiert werden, "[...]

that the product of all the ratios of scale parameters between levels must be identical from the root to all elemental alternatives" (Hensher & Greene, 2002, S. 13).

## 4 Modellschätzung mit der PROC MDC in SAS© 9.1.3

Im Folgenden wird an einem fiktiven Beispiel aus dem Verlagswesen die Schätzung eines Nested Logit Modells mit der *PROC MDC* in SAS© 9.1.3 demonstriert, wobei insbesondere auf die unterschiedlichen Spezifikationen des Nested Logit Modells eingegangen wird. Als Referenz werden die Schätzergebnisse der Kommandos *nlogit* und *nlogitrum* in STATA© 9.1 herangezogen. Die genannten Kommandos in STATA© erlauben eine Validierung der in SAS© 9.1.3 geschätzten Ergebnisse. Das Kommando *nlogit* hat genau wie die Prozedur *PROC MDC* die *non-normalized nested logit* (NNNL) Spezifikation hinterlegt, wohingegen das Kommando *nlogitrum* die mit der Nutzentheorie konsistente *utility maximization nested logit* (RU2 UMNL) Spezifikation schätzt.

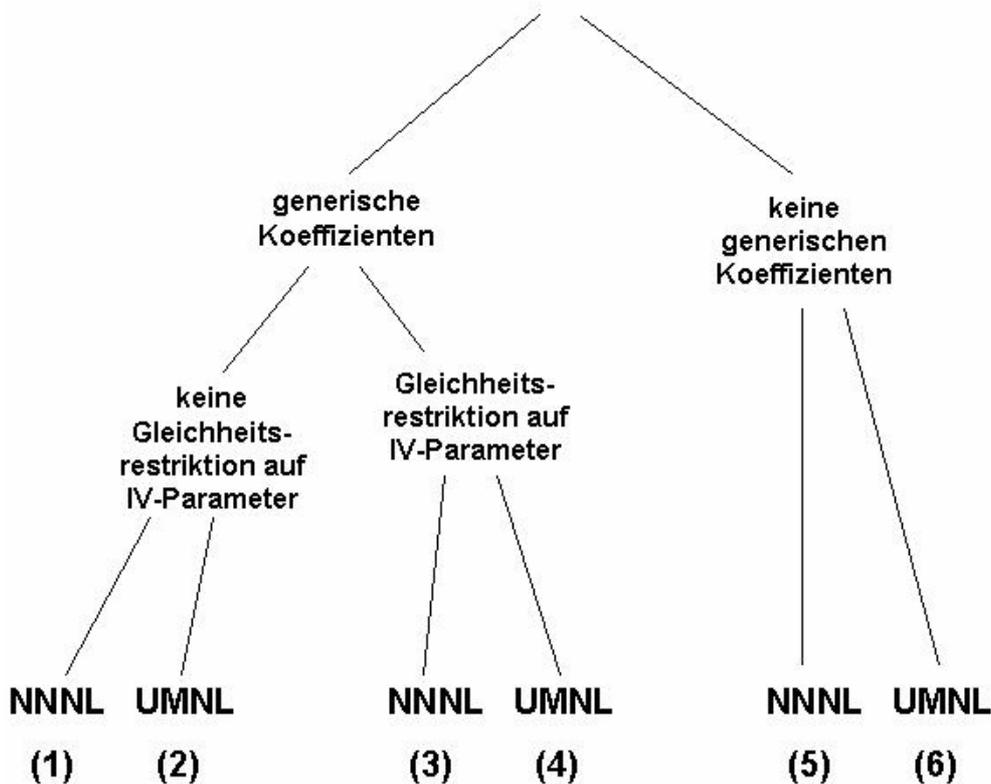
Anhand eines simulierten Datensatzes werden die Auswirkungen der Verwendung unterschiedlicher Spezifikationen des Nested Logit Modells auf die Schätzergebnisse aufgezeigt. Dabei ist von entscheidender Bedeutung, dass der *PROC MDC* in SAS© 9.1.3 die NNNL Form, und somit grundsätzlich kein der Nutzenmaximierungstheorie konsistentes Modell hinterlegt ist.

### 4.1 Modelltypen

Für den Fall, dass nur NNNL-Software zugänglich sein sollte, sind bei der Modellschätzung einige Besonderheiten zu beachten. Dabei ist entscheidend, ob im zu schätzenden Modell ausschließlich **alternativenspezifische** Koeffizienten vorhanden sind, oder aber auch generische Koeffizienten in das Modell einfließen. **Abbildung 2** gibt einen Überblick über die verschiedenen Modelltypen in Kombination mit der verfügbaren Software.

Zunächst ist zu berücksichtigen, ob **generische** Koeffizienten im Modell vorhanden sind. Diese sind für alle Alternativen konstant. Eine Variation des Nutzenbeitrags wird durch die alternativenspezifischen Ausprägungen der zugehörigen Variablen erreicht. Ist das Modell frei von generischen Koeffizienten (Modelle 5 und 6 in **Abbildung 2**), sind die NNNL und die UMNL Spezifikation äquivalent (Heiss, 2002). Die mit der NNNL Software geschätzten Koeffizienten sind jedoch mit dem entsprechenden geschätzten IV-Parameter zu reskalieren. Erst dann ist eine sinnvolle Interpretation möglich. Dabei ist zu beachten, welche Alternative sich in welchem Nest befindet. Zufallsnutzenmaximierende Modelle können beim Auftreten von **generischen** Koeffizienten grundsätzlich **nicht** mit NNNL Software geschätzt werden. Im Gegensatz zur expliziten Skalierung in der UMNL Spezifikation werden die Koeffizienten in der NNNL Spezifikation automatisch und implizit nestspezifisch skaliert. Die im NNNL Modell geschätzten Koeffizienten sind somit nicht die "wahren" Koeffizienten. Zwar

sind die ausgewiesenen Koeffizienten konstant für alle Alternativen, nicht aber die dahinter liegenden "wahren" Koeffizienten. Dies ist eine Verletzung der Definitionsannahme generischer Koeffizienten. Durch die Auferlegung von Restriktionen (Modelle 3 und 4 in **Abbildung 2**) kann jedoch gewährleistet werden, dass auch unter Verwendung von NNNL Software mit der Zufallsnutzentheorie konsistente Koeffizienten geschätzt werden (siehe Abschnitt 3). Diese sind jedoch unbedingt mit dem geschätzten IV-Parameter zu reskalieren. Es darf jedoch nicht vergessen werden, dass jede Restriktion auf die zu schätzenden Parameter mit einem Informationsverlust einhergeht. Befinden sich generische Koeffizienten im Modell und es werden keine Restriktionen auf die Skalenparameter auferlegt (Modelle 1 und 2 in **Abbildung 2**) werden bei Verwendung von NNNL Software falsche Koeffizienten geschätzt. Bei der Modellschätzung mit UMNL Software ist zur Beurteilung der Schätzergebnisse eine Unterscheidung zwischen der RU1 und der RU2 Normalisierung vorzunehmen. Wie in Abschnitt 3 gezeigt wurde, erhält man nur bei der RU2 Normalisierung der UMNL Spezifikation theoriekonsistente und korrekte Schätzergebnisse.



**Abbildung 2:** Übersicht Modelltypen

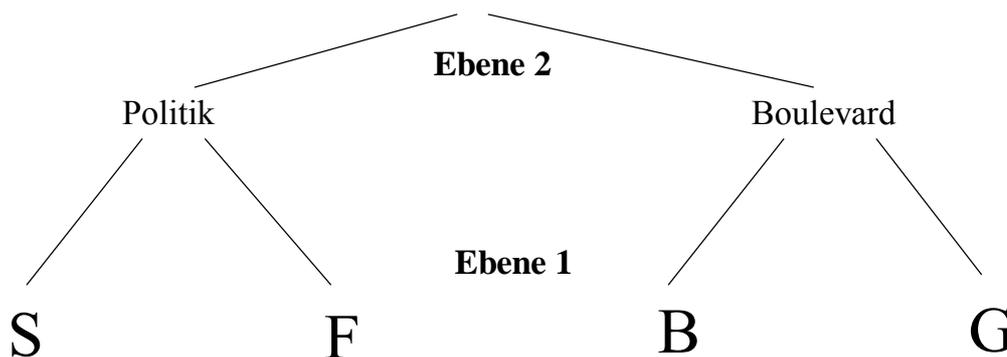
Der Zufallsnutzen  $U_{im}$  jeder Alternative  $im$  ergibt sich aus der Summe einer marginalen Nutzenkomponente  $U_m$  aus Level 2 und einer konditionalen Nutzenkomponente  $U_{i|m}$  aus Level 1, die jeweils aus einem deterministischen Part  $V$  und einem stochastischen Part  $v$  bestehen.

In dieser Studie wird die deterministische marginale Nutzenkomponente  $V_m$  vernachlässigt. Es ist oftmals schwierig, Variablen zu finden, die nest- anstatt alternativenspezifisch sind. Falls doch eine nestspezifische Variable existieren sollte, macht es keinen Unterschied, ob diese Variable für das Nest oder für alle Alternativen innerhalb des Nests spezifiziert wird (Heiss, 2002). Die stochastische marginale Nutzenkomponente  $v_m$ , die alle unbeobachtbaren und vergessenen Faktoren abbildet, ist trotz des Nichtvorhandenseins der deterministischen marginalen Nutzenkomponente  $V_m$  in das Modell zu integrieren. Somit ergibt sich der Gesamtnutzen für die Simulationsstudie aus

$$U_{im} = V_{i|m} + v_{i|m} + v_m .$$

## 4.2 Beispiel

Im Folgenden werden die Kaufwahrscheinlichkeiten von vier Zeitschriften geschätzt, wobei von der in **Abbildung 3** dargestellten Neststruktur ausgegangen wird, die intuitiv gewählt wurde.



**Abbildung 3:** Auswahl von Zeitschriften als Nested Logit Modell

Bei der Simulation von Individualdaten werden der Bildungsstand (EDU), das Alter (AGE) und das Geschlecht (SEX) als haushaltsspezifische exogene Variablen in das Modell einbezogen, sowie der Preis (PRI) der jeweiligen Zeitschrift und etwaige Werbemaßnahmen (PRO) der Verlage als marketingpolitische unabhängige Variablen mit generischen Koeffizienten.

Ferner wird davon ausgegangen, dass jeder der 100 Haushalte eines Panels in jedem der 52 Beobachtungszeitpunkte jeweils einen Kauf vornimmt. Somit erhält man einen Datensatz mit  $n=5.200$  Beobachtungen. Das Alter (AGE=1 mit  $p=0,15$ ; AGE=2 mit  $p=0,20$ ; AGE=3 mit  $p=0,30$ ; AGE=4 mit  $p=0,20$ ; AGE=5 mit  $p=0,15$ ) und der Bildungsstand (EDU=1 mit  $p=0,10$ ; EDU=2 mit  $p=0,30$ ; EDU=3 mit  $p=0,40$ ; EDU=4 mit  $p=0,20$ ) werden als ordinale Variablen, das Geschlecht (männlich, SEX=1 mit  $p=0,48$ ; weiblich, SEX=0 mit  $p=0,52$ ) und die Werbemaßnahmen (Gleichverteilung im Intervall  $[0;1]$  und Rundung) als binäre Variablen und der Preis (Normalverteilung) als

metrische Variable in das Modell integriert. Es wird angenommen, dass die Option des "Outside Good" nicht gegeben ist (d. h. kein "Nicht-Kauf") und zu jedem Beobachtungszeitpunkt nur eine der Alternativen gewählt werden kann.

Typischerweise ist die Variable Preis eines der zentralen Marketing-Mix Elemente bei der Modellierung von Kaufentscheidungen mit Markenwahlmodellen. Um das Beispiel auf andere Branchen übertragbar zu halten, wird bei der hier gewählten Modellierung von Preisschwankungen ausgegangen. Fixierte Preise der einzelnen Zeitschriften, bzw. Preise die nur über einen langen Zeitraum hin schwanken, wie es typischerweise im Verlagswesen zu beobachten ist, werden, neben Gründen der Übertragbarkeit, auch zur Vereinfachung der Schätzung ausgeblendet. Grundsätzlich ist jedoch auch dies bei entsprechender Modellierung in den hier vorgestellten Ansatz integrierbar.

Für die Modellschätzung wird folgende Nutzenfunktion angenommen:

$$V_{im}^h = \eta_i AGE^h + \tau_i SEX^h + \gamma_m EDU^h + \beta_{pri} PRI_i + \beta_{pro} PRO_i$$

Die Alternative  $S$  im Nest *Politik* wird als Referenz deklariert und deren alternativenspezifische Koeffizienten  $\eta_S$  und  $\tau_S$  auf Null gesetzt. Außerdem wird der Koeffizient  $\gamma_{poli}$  als Referenz ebenfalls auf Null gesetzt.

Zur Durchführung der Simulationsstudie wurden zwei Datensätze erzeugt. Dabei wurde die mit der Theorie der Nutzenmaximierung konforme *utility maximization nested logit* Spezifikation mit der Skalenparameter-Normalisierung  $\lambda_m = \lambda_n = 1$  auf Ebene 2 (UMNL RU2) zu Grunde gelegt. Die **Tabellen 3 und 4** unterscheiden sich lediglich in einem Punkt. Während bei der Datensimulation für den Datensatz "RU2 unrestricted" für die beiden Nester „Politik“ und „Boulevard“ verschiedene Skalenparameter auf Ebene 1 angenommen wurden, nämlich  $\mu_{poli} = 1,20$  und  $\mu_{boul} = 1,60$ , wurde bei der Erzeugung des Datensatzes "RU2 restricted" von einem einheitlichen Skalenparameter  $\mu_{poli} = \mu_{boul} = 1,80$  ausgegangen.

### 4.3 Ergebnisse

Die Ergebnisse der Modellschätzungen mit dem Datensatz "RU2 unrestricted" sind in **Tabelle 3** zusammengefasst.

In **Modell 1a** wurde mit *PROC MDC* in SAS© 9.1.3 ein *non-normalized nested logit* (NNNL) Modell geschätzt. Es schätzt eine umgekehrte Beziehung der IV-Parameter in den beiden Nestern. Der IV-Parameter im Nest **Politik** ist kleiner als der IV-Parameter im Nest **Boulevard**, die Korrelation  $(1-(\lambda/\mu)^2)$  der Alternativen innerhalb des Nests **Politik** ist somit größer als im Nest **Boulevard**. Bei der Generierung der simulierten Daten ist jedoch eine stärkere Korrelation im Nest **Boulevard** angenommen worden. **Modell 2a** sollte die Input-Parameter, die bei der Erzeugung des Datensatzes "RU2 unrestricted" verwendet wurden, reproduzieren. Dies gelingt jedoch nicht. Ein zu großer

Einfluss der stochastischen Nutzenkomponente ist als Ursache denkbar und sollte näher untersucht werden.

**Tabelle 3:** Modellschätzung mit Datensatz „RU2 unrestricted“

	Input	SAS <sup>®</sup> PROC MDC NNNL unrestricted <b>1a</b>	STATA <sup>®</sup> nlogitrum RU2 UMNL unrestricted <b>2a</b>	SAS <sup>®</sup> PROC MDC NNNL restricted <b>3a</b>	STATA <sup>®</sup> nlogitrum RU2 UMNL restricted <b>4a</b>
$\eta_F$	-0,50	-0,58***	-0,37***	-0,58***	-0,36***
$\eta_B$	1,00	1,34***	0,90***	1,47***	0,90***
$\eta_G$	1,00	1,33***	0,90***	1,46***	0,90***
$\tau_F$	1,00	1,11***	0,70***	1,11***	0,68***
$\tau_B$	-2,50	-3,16***	-2,18***	-3,55***	-2,19***
$\tau_G$	-3,00	-3,73***	-2,52***	-4,12***	-2,54***
$\gamma_{bowl}$	-1,00	-1,34***	-0,94***	-1,52***	-0,94***
$\beta_{pri}$	-0,50	-0,72***	-0,45***	-0,71***	-0,44***
$\beta_{pro}$	1,00	1,14***	0,70***	1,14***	0,70***
$IV_{poli}$	0,83	0,56***	0,64***	0,62***	0,62***
$IV_{bowl}$	0,63	0,69***	0,60***	0,62***	0,62***
LL		-5.286	-5.286	-5.286	-5.286

$n = 5.200$

\*\*\*  $\alpha = 0,01$ , \*\*  $\alpha = 0,05$ ; \*  $\alpha = 0,10$

Die geschätzten Parameter in **Modell 3a** können durch Multiplikation mit dem für beide Nester gleichgesetzten IV-Parameter in die geschätzten Parameter des **Modells 4a** überführt werden. Bei der Einführung einer IV-Parameter-Restriktion schätzen also das *non-normalized nested logit* (NNNL) und das auf Ebene 2 normalisierte *utility maximization nested logit* (UMNL RU2) Modell sich nur durch einen Skalierungsfaktor unterscheidende Modelle.

Die Ergebnisse der Modellschätzungen mit dem Datensatz "RU2 restricted" sind in **Tabelle 4** zusammengefasst.

**Tabelle 4:** Modellschätzung mit Datensatz „RU2 restricted“

	Input	SAS <sup>©</sup> <i>PROC MDC</i> NNNL unrestricted <b>1b</b>	STATA <sup>©</sup> <i>nlogitrum</i> RU2 UMNL unrestricted <b>2b</b>	SAS <sup>©</sup> <i>PROC MDC</i> NNNL restricted <b>3b</b>	STATA <sup>©</sup> <i>nlogitrum</i> RU2 UMNL restricted <b>4b</b>
$\eta_F$	-0,50	-0,62***	-0,53***	-0,61***	-0,52***
$\eta_B$	1,00	1,16***	1,08***	1,26***	1,08***
$\eta_G$	1,00	1,15***	1,07***	1,25***	1,07***
$\tau_F$	1,00	1,23***	1,07***	1,24***	1,06***
$\tau_B$	-2,50	-2,74***	-2,58***	-3,00***	-2,58***
$\tau_G$	-3,00	-3,36***	-3,11***	-3,63***	-3,11***
$\gamma_{bowl}$	-1,00	-1,13***	-1,07***	-1,25***	-1,07***
$\beta_{pri}$	-0,50	-0,58***	-0,42***	-0,48***	-0,41***
$\beta_{pro}$	1,00	1,15***	0,98***	1,15***	0,98***
$IV_{poli}$	0,56	0,79***	0,87***	0,86***	0,86***
$IV_{bowl}$	0,56	0,94***	0,85***	0,86***	0,86***
LL		-5.019	-5.020	-5.020	-5.020

$n = 5.200$

\*\*\*  $\alpha = 0,01$ , \*\*  $\alpha = 0,05$ ; \*  $\alpha = 0,10$

In **Modell 1b** wurde mit *PROC MDC* in SAS<sup>©</sup> 9.1.3 ein *non-normalized nested logit* (NNNL) Modell geschätzt. Es schätzt in den beiden Nestern einen beträchtlichen Unterschied in den IV-Parametern. Bei der Generierung der simulierten Daten ist jedoch von einheitlichen IV-Parametern und damit Korrelationen innerhalb der Nester ausgegangen worden. Trotz der Schätzung ohne Restriktion in **Modell 2b** schätzt das Modell den Input-Parametern annähernd identische Parameter. Die geschätzten Parameter in **Modell 3b** können durch Multiplikation mit dem für beide Nester

gleichgesetzten IV-Parameter in die geschätzten Parameter des **Modells 4b** überführt werden. **Modell 4b** kann die Input-Werte, die bei der Erzeugung des Datensatzes "RU2 restricted" verwendet wurden, zufriedenstellend reproduzieren.

In allen vier Modellen wird die unterstellte Korrelation innerhalb der beiden Nester nicht reproduziert, sondern in ihrem Ausmaß unterschätzt.

#### 4.4 Fazit

Es konnte gezeigt werden, dass die mit der Prozedur *PROC MDC* in SAS© 9.1.3 geschätzten Koeffizienten eines Nested Logit Modells **nicht** ohne Weiteres interpretiert werden dürfen. Es sind drei Fälle zu unterscheiden: (1) Modell ohne generische Koeffizienten, (2) Modell mit generischen Koeffizienten und mit Gleichheitsrestriktion auf die Skalenparameter, und (3) Modell mit generischen Koeffizienten und ohne Gleichheitsrestriktion auf die Skalenparameter. Im Fall (1) sind die mit der *PROC MDC* geschätzten Koeffizienten mit dem IV-Parameter zu reskalieren, um die "wahren" Koeffizienten zu erhalten. Dies gilt ebenso für die Koeffizienten im Fall (2). Das im Fall (3) geschätzte Modell ist unbrauchbar. Dies ist insofern von besonderer Bedeutung, wenn sich der SAS© Anwender der hier thematisierten Problematik hinsichtlich der existierenden unterschiedlichen Nested Logit Spezifikationen nicht bewusst sein sollte. Die Gefahr einer falschen Modellschätzung wäre dann sehr hoch.

Eine ähnliche Gefahr ist auch bei Verwendung der STATA© Software gegeben. In der Version 9.1 ist nur das Kommando *nlogit* implementiert, das genau wie die Prozedur *PROC MDC* in SAS© 9.1.3 die grundsätzlich nicht theoriekonsistente *non-normalized nested logit* (NNNL) Spezifikation hinterlegt hat. Das Kommando *nlogitrum*, das ein mit der Nutzentheorie konsistentes Modell schätzt, muss erst selbstständig installiert werden. Während *nlogit* darauf besteht, auf jedem Level erklärende Variablen im Modell zu haben, erlaubt *nlogitrum* deren Auftreten lediglich auf der untersten Ebene konditionaler Wahrscheinlichkeiten.

## 5. Zusammenfassung

Obwohl dem Nested Logit Modell aufgrund seiner Fähigkeit, Ähnlichkeiten zwischen Alternativen zuzulassen und durch die teilweise Korrelation der Fehlerterme abzubilden, zunehmend Beachtung geschenkt wird, werden die unterschiedlichen Spezifikationen des Nested Logit Modells kaum differenziert betrachtet. Diese Unterscheidung bekommt aus der Tatsache, dass grundsätzlich nur die RU2 UMNL Form konsistent mit der Zufallsnutzentheorie ist, besondere Relevanz.

Sowohl bei Schätzungen mit realen Daten, als auch bei Simulationsstudien ist zunächst die der zu verwendenden Software hinterlegte Spezifikation zu ermitteln. Wohingegen bei Modellschätzungen mit *utility maximization nested logit* (UMNL) Software keine Besonderheiten zu beachten sind, erweist sich die Schätzung mit *non-normalized nested*

*logit* (NNNL) Software als schwieriger. Nur durch die Auferlegung von Restriktionen auf die IV-Parameter oder die Einführung von Dummy-Nestern können mit der Zufallsnutzentheorie konsistente Schätzergebnisse erzielt werden.

Anhand von simulierten Datensätzen konnte gezeigt werden, dass ohne die Auferlegung von Restriktionen mit NNNL Software kein mit der Zufallsnutzentheorie konsistentes Modell geschätzt werden kann. Nimmt man a priori an, dass sich die Skalenparameter innerhalb der Nester nicht unterscheiden, können die mit NNNL Software (z. B. *PROC MDC* in SAS©) geschätzten Koeffizienten durch Multiplikation mit dem ermittelten IV-Parameter in die mit UMNL Software (z. B. *nlogitrum* in STATA©) geschätzten Koeffizienten überführt werden.

Der SAS© Anwender muss bei der Modellschätzung mit der *PROC MDC* beachten, ob sich generische Koeffizienten im Modell befinden (vgl. **Abbildung 2**). Ist dies nicht der Fall, sind die geschätzten Koeffizienten mit dem entsprechenden nestspezifischen geschätzten IV-Parameter zu reskalieren. Erst dann ist eine sinnvolle Interpretation möglich. Befinden sich generische Koeffizienten im Modell, **muss** auf die Skalenparameter eine Gleichheitsrestriktion auferlegt werden, um mit der Nutzentheorie konsistente und korrekte Schätzergebnisse zu erhalten. Wird im Falle generischer Koeffizienten bei der Modellschätzung keine entsprechende Restriktion eingeführt, sind die Ergebnisse nicht zu gebrauchen.

## Anhang A

### SAS© 9.1.3 Codes

```
proc mdc data=ksfe.zeitschrift_unres outest=ksfe.coeff_unres covout;  
  id id;  
  model decision = edu_boul age_F age_B age_G sex_F sex_B sex_G  
    price promo/  
    type=nlogit  
    choice=(alt 1 2 3 4)  
    maxiter=200;  
  nest level(1) = (1 2 @ 1, 3 4 @ 2),  
  level(2) = (1 2 @ 1);  
  utility u(1, ) = age_F age_B age_G sex_F sex_B sex_G price  
    promo edu_boul;  
run;
```

**Code A1:** Datensatz „RU2 unrestricted“

```
proc mdc data=ksfe.zeitschrift_res outest=ksfe.coeff_res covout;
  id id;
  model decision = edu_boul age_F age_B age_G sex_F sex_B sex_G
    price promo/
    samescale
    type=nlogit
    choice=(alt 1 2 3 4)
    maxiter=200;
  nest level(1) = (1 2 @ 1, 3 4 @ 2),
  level(2) = (1 2 @ 1);
  utility u(1, ) = age_F age_B age_G sex_F sex_B sex_G price
    promo edu_boul;
run;
```

**Code A2:** Datensatz „RU2 restricted“

## Anhang B

### STATA© 9.1 Codes

```
insheet using "/home/till.dannewald/ksfe.zeitschrift_unres.txt"
nlogitgen type= alt(poli: 1 | 2 , boul: 3 | 4)
nlogitrum decision age_f age_b age_g sex_f sex_b sex_g price promo
  edu_boul, group(id) nests(alt type)
```

**Code B1:** Datensatz „RU2 unrestricted“

```
insheet using "/home/till.dannewald/ksfe.zeitschrift_res.txt"
nlogitgen type= alt(poli: 1 | 2 , boul: 3 | 4)
nlogitrum decision age_f age_b age_g sex_f sex_b sex_g price promo
  edu_boul, group(id) nests(alt type) ivc(poli=boul)
```

**Code B2:** Datensatz „RU2 restricted“

### Literatur

- [1] Ben-Akiva, M., Lerman, S. R. (1985): *Discrete Choice Analysis*. Cambridge, The MIT Press
- [2] Carrasco, J. A., de Dios Ortuzar, J. (2002): *Review and assessment of the nested logit model*. *Transport Reviews* 22(2), 197-218

- [3] Chintagunta, P. K., Vilcassim, N. J. (1998): *Empirical implications of unobserved household heterogeneity for manufacturer and retailer pricing*. Journal of Retailing and Consumer Services 5(1), 15-24
- [4] Heiss, F. (2002): *Structural choice analysis with nested logit models*. The Stata Journal 2(3), 227-252
- [5] Hensher, D. A., Greene, W. H. (2002): *Specification and estimation of the nested logit model: alternative normalisations*. Transportation Research B 36(1), 1-17
- [6] Hensher, D. A., Rose, J. M., Greene, W. H. (2005): *Applied Choice Analysis*. Cambridge University Press
- [7] Hunt, G. L. (2000): *Alternative Nested Logit Model Structures and the Special Case of Partial Degeneracy*. Journal of Regional Science 40(1), 89-113
- [8] Koppelman, F. S., Wen, C.-H. (1998a): *Alternative Nested Logit Models: Structure, Properties, and Estimation*. Transportation Research B 32(5), 289-298
- [9] Koppelman, F. S., Wen, C.-H. (1998b): *Nested Logit Models: Which Are You Using?*. Transportation Research Record 1645, 1-7
- [10] McFadden, D. (1978): *Modelling the choice of residential location*. In: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. W. (Hrsg.): *Spatial Interaction Theory and Planning Models*, North-Holland Publishing Company, Amsterdam New York Oxford, 75-96
- [11] McFadden, D. (1981): *Econometric Models of Probabilistic Choice*. In: Manski, C. F., McFadden, D. (Hrsg.): *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge, The MIT Press, 198-272
- [12] Ngobo, P. V. (2005): *Drivers of upward and downward migration: An empirical investigation among theatregoers*. International Journal of Research in Marketing 22, 183-201
- [13] Train, K. E. (2003): *Discrete Choice Methods with Simulation*. Cambridge University Press