Sparsame Modellierung mit logistischen Zufallsnutzenmodellen

Kühnel, Steffen M.

Sparsame Modellierung mit logistischen Zufallsnutzen-modellen

Von Steffen Kühnel

Zusammenfassung

Ein Anwendungsproblem multinomialer Logitmodelle besteht in der die inhaltliche Interpretation erschwerenden hohen Anzahl von Modellparametern bei einer abhängigen Variable mit mehr als zwei Kategorien. Es ist jedoch oft möglich, restriktivere Modelle zu spezifizieren, die weniger Parameter benötigen. In dem Beitrag wird an theoretischen und empirischen Beispielen gezeigt, wie sich solche Restriktionen spezifizieren lassen. Anstelle des multinomialen Logitmodells wird dazu die Verwendung logistischer Zufallsnutzenmodelle vorgeschlagen.

Abstract

A difficulty in applying multinomial logistic regression is the high number of parameters in the statistical model. Using additional restrictions, however, it is often possible to specify more parsimonious models. The paper demonstrates how this can be done by estimating logistic random utility modeis.

In einer Reihe von Beiträgen ist in den ZA-Informationen die Datenanalyse mit logistischen Modellen auf Individualdatenebene vorgestellt und diskutiert worden (vgl. *Ludwig-Mayerhof er,* 1990 und 1992; *Kühnel,* 1990; *Kühnel u.a.,* 1989; ***Urban,*** 1990). Bei einer abhängigen Variable mit mehr als zwei Ausprägungen hat das logistische Regressionsmodell jedoch sehr viele Modellparameter. Die Darstellung von Ergebnissen wird dadurch schnell unübersichtlich. Probleme können auch bei der Interpretation entstehen, ist doch die Signifikanz von Effekten auch davon abhängig, welche Referenzkategorie gewählt ist. *Long* (1987) schlägt als Lösung dieses Problems eine graphische Darstellung und Interpretation vor. Eine Alternative ist, nach Modellen zu suchen, die sparsamer in der Zahl der Modellparameter sind. Wenn die abhängige Variable ordinal ist, können die von *Ludwig-Mayerhofer* (1990; 1992) vorgestellten Modelle eingesetzt werden. Aber auch bei einer nominalskalierten abhängigen Variable lassen sich oft sparsamere Modelle finden. Erleichtert werden solche Spezifikationen, wenn für die Schätzung anstelle des logistischen Regressionsmodells das logistische Zufallsnutzenmodell verwendet wird. Ich möchte in diesem Beitrag Gemeinsamkeiten und Unterschiede des logistischen Regressionsmodells und des logistischen Zufallsnutzenmodells aufzeigen, Beispiele für die Spezifikation von Restriktionen geben und die Vorgehensweise an einem empirischen Anwendungsbeispiel, das der Arbeit von ***Kühnel*** (1993) entnommen ist, verdeutlichen. In einem Anhang stelle ich ein SPSS-Macro vor, mit dem die Parameter logistischer Zufallsnutzenmodelle geschätzt werden können.

1. Das multinomiale Logitmodell und das logistische Zufallsnutzenmodell

Das Kennzeichen einer nominalskalierten Variable ist, daß sich ihre möglichen Ausprägungen nicht in eine eindimensionale Rangordnung bringen lassen. In Erklärungsmodellen für solche Variablen werden daher nicht die Ausprägungen selber, sondern die Auftretenswahrscheinlichkeiten der einzelnen Ausprägungen als Funktion von unabhängigen Variablen aufgefaßt. Im logistischen Regressionsmodell oder multinomialen Logitmodell ist die Wahrscheinlichkeit der Ausprägung oder Kategorie i der abhängigen Variable Y proportional zur Exponentation einer Linearkombination erklärender Variablen X_1, X_2, ..., X_k:

$$\text{(1)} \qquad \pi(Y=i) = \exp(\beta_{i0} + \beta_{i1} \cdot X_1 + \ldots + \beta_{ik} \cdot X_k + \ldots + \beta_{iK} \cdot X_K) \cdot f$$

Mit Hilfe des Proportionalitätsfaktors f wird sichergestellt, daß die Wahrscheinlichkeit zwischen 0 und 1 liegt. Da sich die Wahrscheinlichkeiten aller Kategorien von Y zu 1 summieren müssen, ist f gerade der Kehrwert der Summe der exponentierten Linearkombinationen für alle Kategorien. Wenn i=1,2,..,I für die Kategorien der abhängigen Variable Y steht, gilt somit:[1]

$$\text{(2)} \qquad \pi(Y=i) = \frac{\exp(\beta_{i0} + \ldots + \beta_{ik}X_k + \ldots + \beta_{iK}X_K)}{\sum_{j=1}^{I} \exp(\beta_{j0} + \ldots + \beta_{jk}X_k + \ldots + \beta_{jK}X_K)} \qquad i = 1, 2, \ldots, I$$

Bei der Schätzung der Regressionskoeffizienten β_{ik} des Modells aus Gleichung (2) tritt das Problem auf, daß nicht alle $I \cdot (K+1)$ Regressionskoeffizienten unabhängig voneinander sind, da sich die Wahrscheinlichkeiten aller Kategorien zu 1 summieren müssen und daher bereits I-1 Regressionsgleichungen das Modell vollständig festlegen. Dieses

---

1 Das Modell kann auch so dargestellt werden, daß nicht die Wahrscheinlichkeiten der Kategorien, sondern Wahrscheinlichkeitsverhältnisse oder deren Logarithmen (die sogenannten *Logits)* als Funktionen der erklärenden Variablen aufgefaßt werden (vgl. *Kühnel,* 1990; *Urban,* 1990). Die verschiedenen Formulierungen sind äquivalent und lassen sich ineinander überführen.

Problem wird i.a. dadurch gelöst, daß für eine Referenzkategorie r der abhängigen Variable alle Regressionskoeffizienten a priori auf 0 gesetzt werden. Da die Exponentation von 0 gerade 1 ergibt, ändert sich das Modell zu:

$$
(3) \quad
\begin{aligned}
\pi(Y=i\neq r) &= \frac{\exp(\beta_{i0}+\beta_{i1}X_1+\ldots+\beta_{ik}X_k+\ldots+\beta_{iK}X_K)}{1+\sum_{j=1;j\neq r}^{I}\exp(\beta_{j0}+\beta_{j1}X_1+\ldots+\beta_{jk}X_k+\ldots+\beta_{jK}X_K)} \\
\pi(Y=r) &= \frac{1}{1+\sum_{j=1;j\neq r}^{I}\exp(\beta_{j0}+\beta_{j1}X_1+\ldots+\beta_{jk}X_k+\ldots+\beta_{jK}X_K)}
\end{aligned}
$$

Als Referenzkategorie wird für das multinomiale Logitmodell aus Gleichung (3) üblicherweise die erste oder die letzte Kategorie von Y verwendet.

Abbildung 1: Schematische Darstellung eines Logitmodells

$$
\underline{\pi}(y) = g(\quad \underline{B} \quad \cdot \quad \underline{x} \quad)
$$

$$
\pi\begin{pmatrix} Y=1 \\ Y=2 \\ \ldots \\ Y=I=r \end{pmatrix} = g\left(\begin{pmatrix} \beta_{10} & \beta_{11} & \ldots & \beta_{1K} \\ \beta_{20} & \beta_{21} & \ldots & \beta_{2K} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ X_1 \\ \ldots \\ X_K \end{pmatrix} \right)
$$

Um im folgenden Ähnlichkeiten und Unterschiede zum logistischen Zufallsnutzenmodell aufzuzeigen, ist es nützlich, die Struktur eines multinomialen Logitmodells in Matrizenschreibweise darzustellen. Abbildung 1 zeigt das Schema einer solchen Darstellung. Der Ausdruck "g(...)" steht als abkürzendes Symbol für die nichtlineare Funktion, die die Wahrscheinlichkeiten mit den Linearkombinationen der erklärenden Variablen verknüpft. Die Referenzkategorie ist dadurch gekennzeichnet, daß die ihr zugeordnete Zeile der Matrix B der Regressionskoeffizienten nur Nullen enthält. In der Abbildung ist die letzte Kategorie als Referenzkategorie gewählt.

Im Prinzip läßt sich das multinomiale Logitmodell ähnlich wie ein lineares Regressionsmodell interpretieren. Je größer der Absolutwert eines Regressionsgewichts β_{ik} ist, desto stärker reagiert die Wahrscheinlichkeit $\pi(Y=i)$ auf eine Änderung in der erklärenden Variable X_k. Analog zur linearen Regression erhöht sich bei einem positiven Vorzeichen eher die Chance der betreffenden Kategorie, während sie sich bei einem negativen Vorzeichen eher verringert. Zwei Eigenschaften erschweren jedoch die Interpretation. Zum einen ist die Beziehung zwischen den Wahrscheinlichkeiten der Kategorien von Y und den erklärenden Variablen nichtlinear, und zum anderen beeinflußt eine erklärende Variable X_k die Wahrscheinlichkeiten der Ausprägungen von Y nicht nur über ein, sondern über I-1 Regressionsgewichte β_{ik}.

Stellt man die Beziehung zwischen der Auftretenswahrscheinlichkeit einer Kategorie von Y und den Werten einer erklärenden Variable graphisch dar, ergibt sich eine "S"-Förmige Kurve. Im mittleren Bereich ist die Kurve steiler. Sie flacht an den Rändern um so mehr ab, je stärker sie sich den unteren und oberen Grenzwerten 0 bzw. 1 annähert. Dies bedeutet, daß gleich große Veränderungen in einer erklärenden Variable bei mittleren Ausgangswahrscheinlichkeiten einer Kategorie von Y größere Effekte haben als bei sehr kleinen oder sehr großen Ausgangswahrscheinlichkeiten. Im Unterschied zur linearen Regression kann daher aus einem Regressionskoeffizienten β_{ik} nicht direkt abgelesen werden, um welchen Wert sich die Wahrscheinlichkeit $\pi(Y=i)$ ändert, wenn sich eine erklärende Variable X_k um +1 Einheit ändert. Da die erklärende Variable X_k zudem über weitere Regressionsgewichte auch auf die Wahrscheinlichkeiten anderer Kategorien von Y wirkt, ist es sogar möglich, daß sich bei einem positiven Zuwachs von X_k die Wahrscheinlichkeit $\pi(Y=i)$ verringert, obwohl der Regressionskoeffizient β_{ik} positiv ist.

Die Interpretation eines Logitmodells wird erleichtert, wenn statt Wahrscheinlichkeiten Wahrscheinlichkeitsverhältnisse betrachtet werden. Aus Gleichung (2) bzw. (3) ist ersichtlich, daß für das Wahrscheinlichkeitsverhältnis $\pi(Y=i)/\pi(Y=j)$ gilt:

$$
\begin{aligned}
\text{(4)}\qquad \frac{\pi(Y=i)}{\pi(Y=j)} &= \frac{\exp(\beta_{i0}+\beta_{i1}X_1+\ldots+\beta_{ik}X_k+\ldots+\beta_{iK}X_K)}{\exp(\beta_{j0}+\beta_{j1}X_1+\ldots+\beta_{jk}X_k+\ldots+\beta_{jK}X_K)} \\
&= \exp\big((\beta_{i0}-\beta_{j0})+(\beta_{i1}-\beta_{j1})X_1+\ldots \\
&\qquad +(\beta_{ik}-\beta_{jk})X_k+\ldots+(\beta_{iK}-\beta_{jK})X_K\big)
\end{aligned}
$$

Steigt der Wert einer erklärenden Variable X_k um +1 Einheit an, folgt daher für das Wahrscheinlichkeitsverhältnis zweier Kategorien i und j von Y:

$$
\text{(5)}\qquad \left(\frac{\pi(Y=i)}{\pi(Y=j)}\,|\,X_k=x_k+1\right) = \left(\frac{\pi(Y=i)}{\pi(Y=j)}\,|\,X_k=x_k\right)\cdot\exp(\beta_{ik}-\beta_{jk})
$$

Die Exponentation der Differenz zweier Regressionskoeffizienten β_{ik} **und** β_{jk} einer erklärenden Variable X_k gibt also gerade an, um welchen Faktor sich das Wahrscheinlichkeitsverhältnis der Kategorien i und j der abhängigen Variable Y ändert, wenn die erklärende Variable um +1 Einheit ansteigt. Ist die Differenz der Regressionskoeffizienten positiv, dann ist die Exponentation größer 1 und die relativen Chancen der im Zähler aufgeführten Kategorie i steigen. Das Umgekehrte gilt bei einer negativen Differenz der

Regressionskoeffizienten. Da per definitionem alle Regressionskoeffizienten der Referenzkategorie Null sind, folgt für jedes einzelne Regressionsgewicht β_{ik}, **daß** $\exp(\beta_{ik})$ angibt, wie sich das Wahrscheinlichkeitsverhältnis der Kategorie i zur Referenzkategorie r ändert.

Aufgrund dieser Eigenschaft der Exponentation von Regressionsgewichten bzw. Differenzen von Regressionsgewichten hat ***Long*** (1987) vorgeschlagen, die Exponentation der Differenzen von Regressionskoeffizienten als Maße für die Effekte der erklärenden Variablen zu verwenden. Nachteilig ist allerdings, daß die Anzahl der möglichen Effektkoeffizienten noch größer ist als die ohnehin schon sehr große Zahl von Regressionsgewichten. Die Ergebnisse einer Analyse lassen sich dann kaum mehr übersichtlich darstellen. Die Anzahl der Modellparameter läßt sich jedoch in vielen Fällen aufgrund von theoretischen Überlegungen erheblich reduzieren. So ist bei manchen erklärenden Variablen zu erwarten, daß sie nur einige, aber nicht alle Wahrscheinlichkeitsverhältnisse beeinflussen.

Mit dem logistischen Zufallsnutzenmodell lassen sich solche Vermutungen leicht spezifizieren. Im Unterschied zum logistischen Regressionsmodell geht das logistische Zufallsnutzenmodell davon aus, daß eine erklärende Variable X_{ik} über das Regressionsgewicht β_k nur die Kategorie i der abhängigen Variable Y direkt beeinflußt:

$$\text{(6)} \qquad \pi(Y=i) = \frac{\exp(\beta_1 X_{i1} + \ldots + \beta_k X_{ik} + \ldots + \beta_K X_{iK})}{\sum_{j=1}^{I} \exp(\beta_1 X_{j1} + \ldots + \beta_k X_{jk} + \ldots + \beta_K X_{jK})} \qquad i = 1,2,\ldots,I$$

Der Name "Zufallsnutzenmodell" ist dadurch begründet, daß sich die Modellgleichung (6) aus einer rationalen Handlungstheorie ableiten läßt. Dazu muß angenommen werden, daß die Kategorien der abhängigen Variable Handlungsalternativen sind und die erklärenden Variablen Eigenschaften der Handlungsalternativen beschreiben, die für den subjektiven Nutzen der Handelnden relevant sind. Wenn nun nur die Handlung ausgeführt wird, die den maximalen Nutzen verspricht, die erklärenden Variablen aber nur einen Teil des Nutzens der Alternativen erfassen und weitere nicht gemessene Nutzenkomponenten bestimmten Wahrscheinlichkeitsverteilungen folgen, dann besteht zwischen den realisierten Handlungen und den gemessenen Eigenschaften gerade die in Gleichung (6) beschriebene Beziehung (vgl. *Maier und Weiss,* 1990). Aus dieser Herleitung des Modells folgt selbstverständlich nicht, daß das logistische Zufallsnutzenmodell *nur* dann verwendet werden kann, wenn Handlungen analysiert werden und den Handelnden rationales Verhalten unterstellt werden kann. Anstelle des Wortes "logistisches Zufallsnutzenmodell" wird in der Literatur daher auch der Ausdruck "konditionales Logitmodell" verwendet. Da aber dieser zweite Name nicht ganz eindeutig ist, werde ich nur den ersten Namen für das Modell (6) verwenden.

Die Bedeutung der Modellgleichung (6) wird erleichtert, wenn die Struktur des Modells in Matrizenschreibweise betrachtet wird (vgl. Abbildung 2). Verglichen mit dem Logitmodell aus Abbildung 1 sind im Zufallsnutzenmodell aus Abbildung 2 die Positionen von erklärenden Variablen und Regressionskoeffizienten ausgetauscht. Dieser an sich wenig bedeutsame Umstand erleichtert jedoch die Spezifikation sparsamer Modelle sehr. Wie man nämlich der Modellgleichung (6) bzw. der Abbildung 2 entnehmen kann, wird nun der Einfluß einer erklärenden Variable $\mathbf{X}_{ik}$ nur noch über einen einzigen Regressionskoeffizienten $\boldsymbol{\beta}_k$ vermittelt. Jede erklärende Variable beeinflußt direkt nur noch eine Kategorie der abhängigen Variable. Eine weitere Besonderheit des Modells ist, daß verschiedene erklärende Variablen $\mathbf{X}_{1k}, \mathbf{X}_{2k}, ..., \mathbf{X}_{Ik}$ ein gemeinsames Regressionsgewicht $\boldsymbol{\beta}_k$ haben, über das sie die Wahrscheinlichkeiten der Kategorien von Y beeinflussen.

Abbildung 2: Schematische Darstellung eines logistischen Zufallsnutzenmodells

$$\underset{\sim}{\pi}(y) = g(\underset{\sim}{X} \cdot \underset{\sim}{\beta})$$

$$\pi\begin{pmatrix} Y=1 \\ Y=2 \\ \dots \\ Y=I \end{pmatrix} = g\left(\begin{pmatrix} X_{11} & X_{12} & \dots & X_{1K} \\ X_{21} & X_{22} & \dots & X_{2K} \\ \dots & \dots & \dots & \dots \\ X_{I1} & X_{I2} & \dots & X_{IK} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_K \end{pmatrix}\right)$$

Dem Leser mag aufgefallen sein, daß in Gleichung (6) bzw. Abbildung 2 keine Regressionskonstanten vorkommen. Tatsächlich werden diese wie auch spezifische Effekte einzelner Variablen über "Tricks" spezifiziert. Zur Verdeutlichung zeigt Abbildung 3 ein Beispiel für die Spezifikation eines logistischen Zufallsnutzenmodells, das sowohl Regressionskonstanten enthält, als auch Regressionsgewichte, die jeweils nur für eine einzige erklärende Variable gelten. In der Abbildung 3 haben die Variablen $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$ und A_4 spezifische Regressions gewichte $\boldsymbol{\beta}_{A1}$, $\boldsymbol{\beta}_{A2}$, $\boldsymbol{\beta}_{A3}$ **und** $\boldsymbol{\beta}_{A4}$, die sie *nicht* mit anderen erklärenden Variablen teilen. Dies wird dadurch erreicht, daß in der Matrix der erklärenden Variablen statt weiterer Variablen, die den gleichen Regressionskoeffizienten "benutzen" würden, Konstanten mit dem Wert 0 stehen. Im Kontext des logistischen Zufallsnutzenmodells werden solche Variablen, die mit einem spezifischen Regressionsgewicht eine einzelne Kategorie der abhängigen Variable beeinflussen, "alternativenspezifische Variablen" genannt.

Auf ähnliche Weise lassen sich auch Regressionskonstanten spezifizieren. Anstelle einer alternativenspezifischen Variable, deren Werte über die Fälle variieren, wird hier eine Konstante mit dem Wert 1 eingesetzt. Für diese Pseudovariablen wird auch der Ausdruck "alternativenspezifische Konstanten" verwendet. Wie beim multinomialen Logitmodell, wo bei insgesamt I Kategorien der abhängigen Variable nur I-1 Regressionskonstanten spezifiziert werden können, kann es auch im logistischen Zufallsnutzenmodell nur I-1 alternativenspezifische Konstanten geben.

Abbildung 3: Beispiel für ein logistisches Zufallsnutzenmodell

$$\pi\begin{pmatrix} Y=1 \\ Y=2 \\ Y=3 \\ Y=4 \end{pmatrix} = g\left(\begin{pmatrix} 1 & 0 & 0 & G_1 & A_1 & 0 & 0 & 0 \\ 0 & 1 & 0 & G_2 & 0 & A_2 & 0 & 0 \\ 0 & 0 & 1 & G_3 & 0 & 0 & A_3 & 0 \\ 0 & 0 & 0 & G_4 & 0 & 0 & 0 & A_4 \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{30} \\ \beta_G \\ \beta_{A1} \\ \beta_{A2} \\ \beta_{A3} \\ \beta_{A4} \end{pmatrix}\right)$$

Im Unterschied zu den alternativenspezifischen Variablen ist den Variablen G_1, G_2, G_3 und G_4 ein Regressionsgewicht β_G gemeinsam. Mengen von erklärenden Variablen, denen ein Regressionsgewicht gemeinsam ist, werden auch als "generische Variablen" bezeichnet. Insgesamt zeigt Abbildung 3 also die Struktur eines logistischen Zufallsnutzenmodells, in dem die abhängige Variable Y mit 4 Kategorien durch eine generische Variable G mit den Teilvariablen G_1, G_2, G_3 und G_4 sowie vier alternativenspezifische Variablen A_1, A_2, A_3 und A_4 erklärt wird. Darüber hinaus sind drei alternativenspezifische Konstanten spezifiziert.

Die Interpretation der Regressionsgewichte des logistischen Zufallsnutzenmodells erfolgt analog zur Interpretation der Gewichte des multinomialen Logitmodells. Für die Wahrscheinlichkeitsverhältnisse zweier Kategorien i und j der abhängigen Variable gilt:

$$\text{(7)}\quad \begin{aligned} \frac{\pi(Y=i)}{\pi(Y=j)} &= \frac{\exp(\beta_{i0}1+\beta_1 X_{i1}+\ldots+\beta_k X_{ik}+\ldots+\beta_K X_{iK})}{\exp(\beta_{j0}1+\beta_1 X_{j1}+\ldots+\beta_k X_{jk}+\ldots+\beta_K X_{jK})} \\ &= \exp\big((\beta_{i0}-\beta_{j0})+\beta_1(X_{i1}-X_{j1})+\ldots \\ &\qquad +\beta_k(X_{ik}-X_{jk})+\ldots+\beta_K(X_{iK}-X_{jK})\big) \end{aligned}$$

Aus Gleichung (7) folgt, daß die Erhöhung einer erklärenden Variable X_{ik} um +1 Einheit die Wahrscheinlichkeitsverhältnisse $\pi(Y=i)/\pi(Y=j)$ von zwei verschiedenen Kategorien i und j von Y um den Faktor $\exp(\beta_k)$ ändert:

$$(8) \qquad \left(\frac{\pi(Y=i)}{\pi(Y=j)} \middle| X_{ik}=x_{ik}+1\right) = \left(\frac{\pi(Y=i)}{\pi(Y=j)} \middle| X_{ik}=x_{ik}\right) \cdot \exp(\beta_k) \qquad i \neq j$$

Beeinflußt X_{ik} also nur über ein einziges Regressionsgewicht die abhängige Variable, so ändern sich alle Wahrscheinlichkeitsverhältnisse um den gleichen Faktor.

Abbildung 4: Spezifikation eines Logitmodells als Zufallsnutzenmodell

$$\pi\begin{pmatrix} Y=1 \\ Y=2 \\ Y=3 \\ Y=4 \end{pmatrix} = g(\begin{pmatrix} \beta_{10} & \beta_{11} & \beta_{12} \\ \beta_{20} & \beta_{21} & \beta_{22} \\ \beta_{30} & \beta_{31} & \beta_{32} \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ X_1 \\ X_2 \end{pmatrix})$$

$$= g(\begin{pmatrix} 1 & X_1 & X_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & X_1 & X_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & X_1 & X_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \beta_{10} \\ \beta_{11} \\ \beta_{12} \\ \beta_{20} \\ \beta_{21} \\ \beta_{22} \\ \beta_{30} \\ \beta_{31} \\ \beta_{32} \end{pmatrix})$$

Für empirische Analysen mit dem logistischen Zufallsnutzenmodell ist von Bedeutung, daß sich jedes multinomiale Logitmodell als ein spezielles logistisches Zufallsnutzenmodell spezifizieren läßt. Dazu ist es nur notwendig, die erklärenden Variablen des Regressionsmodells als alternativenspezifische Variablen des Nutzenmodells aufzufassen und den Vektor der erklärenden Variablen für alle Kategorien bis auf die Referenzkategorie zu duplizieren. Abbildung 4 gibt dazu ein einfaches Beispiel, bei dem ein multinomiales

Logitmodell als logistisches Zufallsnutzenmodell dargestellt wird.[1] Das logistische Zufallsnutzenmodell ist daher sehr flexibel. Mit ihm können sowohl die üblichen multinomialen Logitmodelle spezifiziert und geschätzt werden als auch speziellere Modelle, bei denen Koeffizienten auf Null gesetzt werden oder über Kategorien hinweg gleichgesetzt werden. Bevor ich hierzu Beispiele gebe, möchte ich zunächst noch kurz auf die ML-Schätzung der Regressionskoeffizienten, statistische Tests und Zusammenhangsmaße eingehen.

Die bei der linearen Regression übliche einfache Kleinstquadratschätzung ist bei logistischen Modellen nicht optimal. Stattdessen werden die Regressionskoeffizienten in der Regel mit der Maximum-Likelihood-Schätzmethode (ML-Schätzung) ermittelt. Bei dieser Methode werden die Koeffizienten so bestimmt, daß die gemeinsame Wahrscheinlichkeit aller Realisationen der abhängigen Variable maximiert wird. Technisch erfolgt die Berechnung so, daß die mit -2 multiplizierte logarithmierte Likelihood-Funktion minimiert wird. In Anlehnung an die log-lineare Tabellenanalyse möchte ich den Funktionswert dieser Minimierungsfunktion auch als "Devianz" bezeichnen. Als Analogon zum Determinationskoeffizienten R^2 der linearen Regression kann dann als ein Maß für die Erklärungskraft eines Modells die relative Devianzreduktion P^2 berechnet werden. Dieses auch als Pseudo-R^2 *(McFadden, 1914)* bezeichnete Maß gibt an, um wieviel sich die Devianz des Modells gegenüber einem Nullmodell verringert, das nur Regressionskonstanten enthält:

$$P^2 = 1 - \frac{\mathcal{L}_1}{\mathcal{L}_0} \tag{9}$$

In Gleichung (9) steht $\mathcal{L}_1$ für die Devianz des Modells mit allen erklärenden Variablen und $\mathcal{L}_0$ für die Devianz des Konstantenmodells. Durch unterschiedliche Festlegung der Modelle von $\mathcal{L}_0$ und $\mathcal{L}_1$ kann die Formel von P^2 auch für die Definition einfacher und partieller Zusammenhangsmaße verwendet werden. Wenn außer den Regressionskonstanten nur eine einzige erklärende Variable im Modell $\mathcal{L}_1$ ist, gibt P^2 die einfache Erklärungskraft dieser Variable an. Enthält das Nullmodell $\mathcal{L}_0$ bereits erklärende Variablen und wird in $\mathcal{L}_2$ nur eine zusätzliche Variable berücksichtigt, ergibt die Anwendung von Gleichung (9) ein (semi-) partielles Zusammenhangsmaß, das den Anstieg der relativen Devianzreduktion durch diese zusätzliche Variable nach der Berücksichtigung aller Variablen des Nullmodells mißt.

[1] Es gilt übrigens auch umgekehrt, daß jedes logistische Zufallsnutzenmodell als ein multinomiales Logitmodell spezifiziert werden kann. Da es hierzu aber notwendig ist, gezielt einige Regressionskoeffizienten auf Null zu fixieren und andere Koeffizienten miteinander gleichzusetzen, wird dieser Weg in empirischen Anwendungen nicht eingeschritten (vgl. *Kühnel*, 1992; 1993).

Die Signifikanz von P^2 kann über einen Likelihood-Ratio-Test (LR-Test) ermittelt werden. Beim LR-Test werden zwei hierarchisch geschachtelte Modelle gegeneinander getestet. Hierarchisch geschachtelte Modelle unterscheiden sich nur dadurch, daß das strengere Modell Restriktionen über Regressionskoeffizienten des weniger strengen Modells postuliert. Treffen die Restriktionen des strengeren Modells in der Grundgesamtheit zu, dann ist die zweifache Differenz der logarithmierten Likelihood-Funktion der beiden Modelle asymptotisch chiquadrat-verteilt, wobei sich die Freiheitsgrade aus der Anzahl der Restriktionen ergeben. Beim Test von P^2 ist das strengere Modell das Nullmodell. Es postuliert, daß alle zusätzlichen Regressionsgewichte im Modell $\mathcal{L}_1$, die nicht bereits im Nullmodell spezifiziert worden sind, in der Grundgesamtheit Null sind. Ist diese Hypothese richtig, dann ist die Differenz der Devianzen chiquadrat-verteilt. Die Anzahl der Freiheitsgrade ist hier die Differenz der Regressionskoeffizienten in den beiden Modellen.

Neben den LR-Tests können auch Wald-Tests angewendet werden. Der Name dieser Tests erinnert an den Statistiker *A. Wald.* Wald-Tests beruhen darauf, daß die ML-Schätzungen der Regressionskoeffizienten asymptotisch um die Populationsparameter multinormalverteilt sind. Dann sind auch beliebige Linearkombinationen der Regressionskoeffizienten, die sogenannten "Kontraste", multinormalverteilt. Da die quadratische Form einer Multinormalverteilung chiquadrat-verteilt ist, lassen sich Kontraste simultan über die Chiquadrat-Verteilung testen. Ein einzelner Kontrast kann auch über die Standardnormalverteilung getestet werden. Ein Beispiel für einen Wald-Test ist der bekannte T-Test der Signifikanz eines einzelnen Regressionskoeffizienten.[1]

2. Beispiele für die Spezifikation von Restriktionen

Aufgrund ihrer engen Verwandtschaft gelten alle Aussagen über die ML-Schätzung, die Verwendung der relativen Devianzreduktion und inferenzstatistische Tests sowohl für das multinomiale Logitmodell wie für das logistische Zufallsnutzenmodell. Welches Modell angewendet wird, spielt im Prinzip keine Rolle. Erst wenn Restriktionen zur Einsparung von Modellparametern spezifiziert werden sollen, bietet das Zufallsnutzenmodell Vorteile. Ich möchte dies an einem Beispiel zeigen, bei dem die Variation einer abhängigen Variable Y mit vier Ausprägungen durch drei Prädiktoren X_1, X_2 und X_3 erklärt werden soll. Wird der Zusammenhang durch ein multinomiales Logitmodell spezifiziert, sind insgesamt 12 Regressionskoeffizienten zu schätzen (vgl. Abbildung 5).

[1] Für Einzelheiten vgl. *Kühnel* (1993).

Abbildung 5: Beispiel eines Modells mit maximaler Parameterzahl

$$\begin{pmatrix} \pi(Y=1) \\ \pi(Y=2) \\ \pi(Y=3) \\ \pi(Y=4) \end{pmatrix} = g(\begin{pmatrix} \beta_{10} & \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{20} & \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{30} & \beta_{31} & \beta_{32} & \beta_{33} \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ X_1 \\ X_2 \\ X_3 \end{pmatrix})$$

Die Zahl der Regressionsgewichte läßt sich deutlich reduzieren, wenn zusätzliche Annahmen über die Wirkungsbeziehung möglich sind. Angenommen, es werde vermutet, daß die erklärende Variable X_1 nur die Chancen der zweiten Kategorie von Y beeinflußt, die relativen Chancen der übrigen Kategorien dagegen nicht tangiert. Weiter soll angenommen werden, daß X_2 nicht das Wahrscheinlichkeitsverhältnis zwischen der ersten und vierten Kategorie von Y beeinflußt. Schließlich soll davon ausgegangen werden, daß X_3 in jeweils gleichem Maße die Chancen der ersten und vierten Kategorie sowie der zweiten und dritten Kategorie von Y beeinflußt. Ein logistisches Zufallsnutzenmodell, das diesen Vermutungen Rechnung trägt, benötigt nur noch 7 Regressionskoeffizienten (vgl. Abbildung 6).

Abbildung 6: Spezifikation des restriktiven Modells als Zufallsnutzenmodell

$$\begin{pmatrix} \pi(Y=1) \\ \pi(Y=2) \\ \pi(Y=3) \\ \pi(Y=4) \end{pmatrix} = g(\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & X_3 \\ 0 & 1 & 0 & X_1 & X_2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & X_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & X_3 \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{30} \\ \beta_{21} \\ \beta_{22} \\ \beta_{32} \\ \beta_{13}=\beta_{33} \end{pmatrix})$$

Wendet man die Gleichung (8) auf das in Abbildung 6 wiedergegebene Modell an, wird deutlich, daß eine Erhöhung der Variable X_1 um +1 Einheit die Wahrscheinlichkeiten der relativen Chancen der zweiten Kategorie zu allen anderen Kategorien um $\exp(\beta_{21})$ verändert, während die Wahrscheinlichkeitsverhältnisse der übrigen Kategorien unverändert bleiben. Verändert sich X_2 um +1 Einheit, ändern sich die Chancen von Kategorie 2 von Y gegenüber den Kategorien 1 und 4 jeweils um den Faktor $\exp(\beta_{22})$. Gleichzeitig ändern sich die Chancen der Kategorie 3 gegenüber den Kategorien 1 und 4 um den Faktor $\exp(\beta_{32})$. Das Chancenverhältnis der Kategorien 2 zur Kategorie 3 ändert sich dann nach Gleichung (5) um den Faktor $\exp(\beta_{22}-\beta_{32})$. Das Wahrscheinlichkeitsverhältnis der Kategorien 1 und 4 wird - wie oben unterstellt - von einer Änderung der Variable X_2 nicht betroffen. Eine Änderung von X_3 um +1 Einheit würde schließlich das Chancenverhältnis von Kategorie 1 zu Kategorie 2 um den Faktor $\exp(\beta_{13})$ verändern. Um den gleichen Faktor würden sich auch die Verhältnisse der Kategorie 1 zu Kategorie 3, der Kategorie 4 zu Kategorie 2 und der Kategorie 4 zu Kategorie 3 ändern. Die relativen Chancen von Kategorie 1 zu Kategorie 4 und von Kategorie 2 zu Kategorie 3 blieben dagegen unverändert.

Denkbar sind auch komplexere Restriktionen. So könnte etwa als eine weitere Restriktion angenommen werden, daß der Einfluß der Variable X_1 auf die Kategorie 2 von Y doppelt so groß ist wie der Einfluß der Variable X_2 auf diese Kategorie. Die Spezifikation dieser zusätzlichen Restriktion ist in Abbildung 7 wiedergegeben.

Abbildung 7: Ein noch restriktiveres Modell

$$\begin{pmatrix} \pi(Y=1) \\ \pi(Y=2) \\ \pi(Y=3) \\ \pi(Y=4) \end{pmatrix} = g\left(\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & X_3 \\ 0 & 1 & 0 & 2\cdot X_1+X_2 & 0 & 0 \\ 0 & 1 & 0 & 0 & X_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & X_3 \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{30} \\ \frac{1}{2}\beta_{21}=\beta_{22} \\ \beta_{32} \\ \beta_{13}=\beta_{33} \end{pmatrix} \right)$$

Zu beachten ist hierbei, was es bedeutet, wenn der Regressionskoeffizient β_{22} doppelt so groß ist wie der Koeffizient β_{21}. Wenn X_1 um +1 Einheit ansteigt, dann ändern sich die Wahrscheinlichkeitsverhältnisse von Kategorie 2 von Y zu den anderen Kategorien um den Faktor $\exp(2\beta_{22})=(\exp(\beta_{22}))^2$. Ein doppelt so großer Regressionskoeffizient bewirkt also bezogen auf die Wahrscheinlichkeitsverhältnisse eine Quadrierung des Änderungsfaktors. Soll sich dagegen das Wahrscheinlichkeitsverhältnis verdoppeln, dann hätte die Variable X_2 nicht mit 2, sondern mit dem natürlichen Logarithmus von 2 (=0.693) multipliziert werden müssen.

3. Ein Anwendungsbeispiel: der Zusammenhang zwischen Teilnahmeabsicht und Teilnahmeverhalten bei der Volkszählung 1987

Die Nutzung der sparsamen Modellierung mit Hilfe des logistischen Zufallsnutzenmodells setzt voraus, daß theoretisch oder empirisch begründete Hypothesen über Modell-restriktionen vorliegen. Daß diese Bedingung auch erfüllbar ist, möchte ich an einem empirischen Beispiel zeigen. Das Beispiel ist einer umfangreicheren Arbeit entnommen, die im Januar 1993 erscheinen wird ***(Kühnel,*** 1993). In dieser Arbeit habe ich versucht, Teilnahmeabsicht und Teilnahmeverhalten bei der Volkszählung 1987 zu erklären. Als Erklärungsansatz wird die *Theorie geplanter Handlungen* von *Ajzen* (1988) herangezogen, die eine Weiterentwicklung der *Theorie bedachter Handlungen* von *Fishbein und Ajzen* (1975; *Ajzen und Fishbein,* 1980) ist. Nach dieser Theorie soll die Erklärung von Handlungen über die Erklärung der Handlungsabsicht erfolgen, da die Handlungsabsicht die entscheidende Determinante einer Handlung sei. Unterstellt wird hierbei ein enger Zusammenhang zwischen Handlungsabsicht und Handlungsausführung. Um im folgenden das Anwendungsbeispiel möglichst einfach zu halten, möchte ich mich nur auf den bivariaten Zusammenhang zwischen diesen beiden Variablen konzentrieren.

Datenbasis für die empirische Analyse sind die ersten beiden Panelwellen der Kölner Begleituntersuchung zur Volkszählung. In der ersten Welle vor der Volkszählung wurde das beabsichtigte Antwortverhalten bei der Zählung und in der zweiten Welle nach der Durchführung der Volkszählung das tatsächliche Antwortverhalten erfragt.[1] Sowohl für das beabsichtigte wie für das berichtete Antwortverhalten gab es vier Antwortmöglichkeiten: vollständiges und ehrliches Beantworten aller Fragen *(Kooperation),* die Nichtbeantwortung von Fragen *(Unvollständigkeit),* Täuschung durch bewußte Falschangaben *(falsche Angaben)* und die Verweigerung jeglicher Teilnahme *(Boykott).*[2] Tabelle 1 zeigt die Kreuztabellierung des berichteten gegen das beabsichtigte Antwortverhalten.

Da sowohl die erklärende wie die abhängige Variable nominalskaliert sind, ist für die Analyse die erklärende Variable (Antwortabsicht) in vier 0/1-kodierte Dummy-Variablen zerlegt, die für jeweils eine Antwortalternative stehen.[3] Drei Dummy-Variablen determinieren dabei den Wert des vierten Dummy. Analog zur Vorgehensweise bei der linearen Regressions- bzw. Varianzanalyse werden daher auch in einem multinomialen Logitmodell nur maximal drei Dummy-Variablen als erklärende Variablen eingesetzt. Das De-

---

[1] Die Daten der Kölner Begleituntersuchung sind unter den Studiennummern 1583 bis 1590 in den Bestand des Zentralarchivs aufgenommen. Angaben zum Erhebungsdesign finden sich bei *Scheuch u.a.* (1989) und ***Kühnel*** (1993).

[2] Der Wortlaut der Fragen ist in *Kühnel* (1993: 107) wiedergegeben. Dort findet sich auch eine Diskussion über die Angemessenheit der Operationalisierung für eine Anwendung im Sinne der *Theorie geplanter Handlungen.*

[3] Da alle Variablen des Modells nominalskaliert sind, könnte die Analyse hier auch mit Hilfe von log-linearen Modellen erfolgen. Tatsächlich würden sich bei entsprechender Modellierung die gleichen Modellparameter ergeben. Unterschiede gäbe es bei der Devianz, die sich bei log-linearen Modellen auf Zellenhäufigkeiten bezieht und bei Logitmodellen auf Individualebene auf die Wahrscheinlichkeiten jedes einzelnen Falls.

Tabelle 1: Zusammenhang zwischen beabsichtigtem und berichtetem Antwortverhalten bei der Volkszählung 1987

Berichtetes Teilnahmeverhalten (nach VZ)	*Beabsichtigtes Teilnahmeverhalten (vor VZ)*			
	Kooperation	*Unvollständigkeit*	*Falsche Angaben*	*Boykott*
Kooperation	97.4 %	89.6 %	77.4 %	51.4 %
Unvollständigkeit	1.8 %	4.9 %	9.7 %	14.3 %
Falsche Angaben	0.5 %	4.5 %	11.3 %	5.7 %
Boykott	0.3 %	1.1 %	1.6 %	28.6 %
(Anzahl)	(653)	(268)	(62)	(35)

Quelle: *Kühnel*, 1993, S. 191

sign eines solchen Modells ist mit dem in Abbildung 5 wiedergegebenen Modell identisch, wobei Y hier für das berichtete Antwortverhalten steht und X_1 bis X_3 die drei 0/1-kodierten Dummy-Variablen für die Kategorien "Kooperation", "Unvollständigkeit" und "falsche Angaben" sind.

Abbildung 8: Erklärung des Antwortverhaltens durch die Antwortabsicht über alternativenspezifische Variablen

$$\begin{pmatrix} \pi\,(\textit{Kooperation}) \\ \pi\,(\textit{Unvollständigkeit}) \\ \pi\,(\textit{falsche Angaben}) \\ \pi\,(\textit{Boykott}) \end{pmatrix} = g\left(\begin{pmatrix} 1 & 0 & 0 & D_{Koop.} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & D_{Unv.} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & D_{falsch} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & D_{Boyk.} \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{30} \\ \beta_{Koop.} \\ \beta_{Unv.} \\ \beta_{falsch} \\ \beta_{Boyk.} \end{pmatrix} \right)$$

Anstelle des multinomialen Logitmodells liegt eine alternative Spezifikation nahe, nach der jede Kategorie des berichteten Antwortverhaltens nur durch die Dummy-Variable beeinflußt wird, die für die entsprechende Kategorie des beabsichtigten Antwortverhaltens steht. Diese Spezifikation läßt sich durch ein logistisches Zufallsnutzenmodell mit vier alternativenspezifischen Variablen realisieren. Die Anzahl der Modellparameter reduziert sich hier von 12 auf 7. Wird angenommen, daß die Effekte aller Dummy-Variablen gleich hoch sind, dann bilden die vier Dummy-Variablen eine generische Variable. Dieses Modell hat neben den drei Regressionskonstanten nur noch ein einziges Regressionsgewicht.

Abbildung 9: Erklärung des Antwortverhaltens durch die Antwortabsicht über generische Variablen

$$\begin{pmatrix} \pi\,(\mathit{Kooperation}) \\ \pi\,(\mathit{Unvollständigkeit}) \\ \pi\,(\mathit{falsche\ Angaben}) \\ \pi\,(\mathit{Boykott}) \end{pmatrix} = g(\begin{pmatrix} 1 & 0 & 0 & D_{Koop.} \\ 0 & 1 & 0 & D_{Unv.} \\ 0 & 0 & 1 & D_{falsch} \\ 0 & 0 & 0 & D_{Boyk.} \end{pmatrix} \begin{pmatrix} \beta_{10} \\ \beta_{20} \\ \beta_{30} \\ \beta_{ABSICHT} \end{pmatrix})$$

Abbildung 8 zeigt die Spezifikation über alternativenspezifische Variablen, Abbildung 9 die Spezifikation über eine generische Variable. In den Abbildungen steht "$D_{Koop.}$", "$D_{Unv.}$", "D_{falsch}" und "$D_{Boyk.}$" für die Dummy-Variablen der Handlungsabsichten Kooperation, unvollständige Angaben, falsche Angaben und Boykott.

Tabelle 2: Erklärungskraft der drei Modelle (n: 1018)

Modell	P^2	χ^2	df	Prob.
Multinomiales Logitmodell	15.8 %	110.2	8	< .001
Modell mit alternativenspezifischen Variablen	15.3 %	107.2	4	< .001
Modell mit generischen Variablen	8.2 %	57.2	1	< .001

In Tabelle 2 sind die relativen Devianzreduktionen **(P^2)** der drei Modelle und deren Signifikanzen wiedergegeben. Interessant ist zunächst, daß die Erklärungskraft des Modells mit alternativenspezifischen Variablen praktisch genau so hoch ist wie die Erklärungskraft des multinomialen Logitmodells. Die Erklärungskraft des Modells mit einer generischen Variable ist dagegen deutlich geringer. Dieses letzte Modell ist in das Modell mit alternativenspezifischen Variablen hierarchisch geschachtelt. Der Unterschied zwischen den beiden Modellen besteht allein darin, daß das Modell mit der generischen Variable postuliert, daß die vier Regressionsgewichte des Modells mit alternativenspezifischen Variablen gleiche Werte aufweisen. Die beiden Modelle lassen sich somit gegeneinander testen. Für den Test müssen die Chiquadratwerte der relativen Devianzreduktionen der beiden Modelle voneinander abgezogen werden. Es ergibt sich ein Wert von 50.0. Die Differenz der Modellparameter beträgt 3. Die Nullhypothese, daß alle Regressionsgewichte gleich groß sind, ist daher selbst bei einem Signifikanzniveau von 1% abzuweisen, da der Chiquadratwert von 50.0 deutlich größer ist als der kritische Wert

11.34 einer Chiquadrat-Verteilung mit drei Freiheitsgraden. Zum gleichen Ergebnis kommt der Wald-Test dieser Hypothese, der einen Chiquadratwert von 48.1 ergibt

Ein analoger Test für die beiden ersten Modelle ist nicht möglich, da das Modell mit alternativenspezifischen Variablen nicht hierarchisch in das Logitmodell geschachtelt ist. Gleichwohl spricht die größere Sparsamkeit des Modells mit alternativenspezifischen Variablen bei nur unmerklich verringerter Erklärungskraft für die Bevorzugung dieses Modells vor dem Logitmodell. Hinzu kommt die leichtere Interpretierbarkeit, da in dem Modell eine Dummy-Variable nur noch über ein Regressionsgewicht Einfluß auf die Wahrscheinlichkeiten der Handlungsalternativen hat.

Tabelle 3: Geschätzte Regressionsgewichte des logistischen Zufallsnutzenmodells mit alternativenspezifischen Variablen

Modellparameter	Gewicht (β)	T-Wert	Effekt (e^{β})
Kooperation	2.158	6.5	8.65
Unvollständigkeit	-1.100	-2.7	1 : 3.00
Falsche Angaben	1.079	2.3	2.94
Boykott	3.369	6.0	29.06

In Tabelle 3 sind die Ergebnisse der Parameterschätzung wiedergegeben. Die geschätzten Regressionsgewichte der vier Dummy-Variablen sind in der Tat sehr verschieden. Den höchsten Effekt hat die Boykottabsicht. Hat ein Befragter vor der Volkszählung angegeben, daß er boykottieren will, dann erhöht dies die tatsächlichen Boykottchancen relativ zu allen anderen Verhaltensalternativen um etwa das Neunundzwanzigfache. Wurde bei der Frage nach dem beabsichtigten Teilnahmeverhalten die Alternative "Kooperation" angegeben, erhöht dies die relativen Chancen tatsächlicher Kooperation um den Faktor 8.6. Bei denjenigen, die "falsche Angaben" machen wollten, steigt die relative Chance der Realisierung dieser Verhaltensmöglichkeit um nicht ganz das Dreifache. Überraschend ist der Effekt bei der Kategorie "Unvollständigkeit". Hier sinken die relativen Chancen unvollständiger Antworten um das Dreifache, wenn ein Befragter in der ersten Welle angegeben hat, unvollständige Angaben machen zu wollen. Bei dieser Verhaltens alternative besteht also ein negativer Effekt zwischen Verhaltensabsicht und Verhaltensausführung. Betrachtet man daraufhin noch einmal Tabelle 1, ist dieses Ergebnis nicht unplausibel: Der Anteil derjenigen, die über unvollständige Angaben bei der Volkszählung berichten, ist bei Befragten, die ursprünglich boykottieren oder falsche Angaben machen wollten, größer als bei denjenigen, die bereits von vornherein ihre Angaben unvollständig geben wollten.

Die Verschiedenheit der Effekte erklärt, warum es nicht möglich ist, Handlungsabsicht und berichtetes Verhalten nur über ein einziges Regressionsgewicht zu verknüpfen. Daß große Unterschiede in den Effekten allerdings nicht notwendigerweise auch statistisch signifikant sind, zeigt der Test auf Gleichheit der Effekte von "Kooperation" und "Boykott". Der LR-Test führt hier nur zu einem Chiquadratwert von 3.336, was bei einem Freiheitsgrad einem Wahrscheinlichkeitswert von 6.8 % entspricht. Zu fast dem gleichen Wert führt der Wald-Test mit einem Chiquadratwert von 3.235 (Prob.: 7.2 %). Bei einer α-Fehlerwahrscheinlichkeit von 5 % kann die Nullhypothese der Gleichheit der beiden Effekte also nicht verworfen werden, obwohl sich die Regressionskoeffizienten und Effekte deutlich unterscheiden. Dies ist eine Folge der geringen Anzahl von nur 35 Befragten, die von Boykott berichten. Bei sehr schiefen Verteilungen sollte allerdings überlegt werden, das Signifikanzniveau zur Verminderung der Risiken von Fehlern zweiter Art (falsche Beibehaltung einer Nullhypothese) auf 10 % zu setzen.

Nach der methodisch-statistischen Diskussion der Ergebnisse möchte ich noch kurz darauf eingehen, welche inhaltlichen Schlußfolgerungen zu ziehen sind. Mit einer relativen Devianzreduktion (P^2) von nur 15 % besteht ein eher mäßiger Zusammenhang zwischen der in der ersten Panelwelle erhobenen Verhaltensabsicht und dem in der zweiten Welle berichteten Antwortverhalten bei der Volkszählung. Darüber hinaus ist die Beziehung bei der Verhaltensalternative "Unvollständigkeit" negativ. Die theoretischen Vorstellungen von *Fishbein* und *Ajzen*, nach denen die Handlungsabsicht unmittelbare und (neben der nicht erfaßten objektiven Kontrolle über das Verhalten) einzige direkte Determinante des Verhaltens ist, werden hierdurch offenbar nicht gestützt. Der geringe Zusammenhang zwischen Handlungsabsicht und -realisation kann allerdings auch Folge von Änderungen sein, die zwischen den Erhebungszeitpunkten eingetreten sind. Daß solche Änderungen tatsächlich eingetreten sind und sich bei Berücksichtigung dieser Änderungen ein ganz anderes Bild bietet, zeigen weitere Analysen der Kölner Begleituntersuchung zur Volkszählung, auf die ich hier aus Platzgründen nicht eingehen kann (vgl. *Kühnel*, 1993).

4. Diskussion

Ich habe in diesem Beitrag zu zeigen versucht, daß sich in multinomialen logistischen Modellen durch Berücksichtigung von Restriktionen die hohe Zahl der Modellparameter reduzieren läßt. Die Interpretation eines Modells wird dadurch sehr erleichtert. Die Regressionskoeffizienten solcher restriktiver Modelle lassen sich leichter schätzen, wenn anstelle des multinomialen Logitmodells das logistische Zufallsnutzenmodell angewendet wird. Leider ist das logistische Zufallsnutzenmodell zur Zeit nicht als Standardprozedur in den Programmsystemen BMDP, SAS oder SPSS verfügbar. Die Modelle lassen sich jedoch über das Zusatzmodul "LOGIT" im Programmsystem SYSTAT schätzen ***(Steinberg und Colla,*** 1991). Das Modell steht auch in ökonometrischen Programmsystemen wie LIMDEP ***(Greene,*** 1990) und in Spezialprogrammen wie KALOS *(Röding u.a.,* 1985) zur Verfügung.[1] Die Modellschätzungen für diesen Beitrag habe ich mit einem eigenen SPSS-Macro berechnet. Eine Beschreibung des Macros, das auf Anforderung Interessierten zur Verfügung gestellt wird, ist im Anhang wiedergegeben.

Die hier vorgeschlagenen restriktiven logistischen Modelle setzen neben der Verfügbarkeit von Computerprogrammen zur Schätzung der Regressionskoeffizienten auch voraus, daß sich in Anwendungen Restriktionen finden und empirisch halten lassen. Denkbar wäre es, einfach so lange verschiedene Modelle auszuprobieren, bis ein sparsames Modell gefunden ist, das gegenüber dem multinomialen Logitmodell ohne Restriktionen keine oder nur geringe Einbußen an Erklärungskraft hat. Solches "data fitting" scheint mir jedoch aus zwei Gründen nicht empfehlenswert zu sein. Zum einen gibt es bereits bei nur wenigen erklärenden Variablen eine Vielzahl möglicher Modelle, die bei vertretbarem Zeitaufwand kaum oder gar nicht alle geschätzt werden können. Darüber hinaus besteht bei dieser Vorgehensweise die Gefahr, ein Modell zu finden, das die Daten einer Stichprobe zwar sehr sparsam beschreibt, an anderen Datensätzen jedoch versagt. Vielversprechender ist es meiner Ansicht nach, über theoretische Überlegungen zu testbaren Restriktionen zu gelangen. Notwendig ist dazu die intensive Auseinandersetzung mit dem inhaltlichen Forschungsproblem. Auch im Hinblick auf die Kommunikation zwischen empirischen Sozialforschern und sozialwissenschaftlichen Theoretikern scheint mir dies durchaus wünschenswert zu sein.

Literatur

Ajzen, I. (1988), Attitudes, Personality, and Behavior. Milton Keynes: Open University Press.

Ajzen, I. und M. Fishbein (1980), Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs, N.J.: Prentice Hall.

Fishbein, M. und I. Ajzen (1975), Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Reading: Addison-Wesley.

Greene, W.H. (1990), LIMDEP Version 5.1. New York: Econometric Software Inc.

[1] Die PC-Version des Programms KALOS wird neben logistischen Modellen zur Panelanalyse auch in der neuesten Version des Programmsystems TDA von *Götz Rohwer* zur Verfügung stehen, die voraussichtlich im Januar 1993 ausgeliefert wird.

Kühnel, *S.-M.* (1990), Lassen sich mit SPSS^x-Matrix anwenderspezifische Analyseprobleme lösen? Ein Anwendungstest am Beispiel der multinomialen logistischen Regression. ZA-Information 27, S. 89-109.

Kühnel, S.-M. (1992), Some Remarks on the Use of Multinomial Logit Models in Sociological Research. Köln: Manuskript.

Kühnel, S.-M. (1993), Zwischen Boykott und Kooperation. Teilnahmeabsicht und Teilnahmeverhalten bei der Volkszählung 1987 (Empirische und methodologische Beiträge zur Sozialwissenschaft, Bd. 11). Frankfurt u.a.: P. Lang (erscheint im Januar 1993).

Kühnel, *S.-M.,* ***W.*** *Jagodzinski, M. Terwey* (1989), Teilnehmen oder Boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSS*. ZA-Information 25, S. 44-75.

Long, J.S. (1987), A Graphical Method for the Interpretation of Multinomial Logit Analysis. Sociological Methods and Research, Jg. 15, S. 420-446.

Ludwig-Mayerhofer, W. (1990), Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen. ZA-Information 27, S. 62-88.

Ludwig-Mayerhofer, W. (1992), Statistik-Software zur Schätzung von Regressionsmodellen für ordinale abhängige Variablen. ZA-Information 31 (in diesem Heft).

Maier, G. und P. Weiss (1990), Modelle diskreter Entscheidungen - Theorie und Anwendungen in den Sozial- und Wirtschaftswissenschaften. Wien u.a.: Springer.

McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior. In: *P. Zarembka* (Hrsg.), Frontiers in Econometrics. New York: Academic Press.

Röding, M., U. Küsters und G. Arminger (1985), KALOS Version 1.0. Ein interaktives Programmsystem zur Analyse kategorialer Logitmodelle. Programmbeschreibung. Köln: Zentralarchiv für empirische Sozialforschung.

Scheuch, E.K., L. Gräf und S.-M. Kühnel (1989), Volkszählung, Volkszählungsprotest und Bürgerverhalten. Ergebnisse der Begleituntersuchung zur Volkszählung 1987. Stuttgart: Metzler-Poeschel.

Steinberg, D. und Ph. Colla (1991), Logit: A supplementary module for SYSTAT. Evanston, IL: SYSTAT Inc.

Urban, D. (1990), Multinomiale LOGIT-Modelle zur Bestimmung der Abhängigkeitsstruktur qualitativer Variablen mit mehr als zwei Ausprägungen. ZA-Information 26, S. 36-61.

Anhang: Dokumentation des SPSS^x-Macros UTIL

Aufbauend auf eine frühere Arbeit *(Kühnel,* 1990), habe ich zur Schätzung der Modellparameter des logistischen Zufallsnutzenmodells ein Macro geschrieben, das die Parameter des logistischen Entscheidungsmodells nach der ML-Methode schätzt. In der Modellspezifikation können sowohl alternativenspezifische als auch generische Variablen spezifiziert werden. Alternativenspezifische Konstanten werden als alternativenspezifische Variablen spezifiziert. Sie müssen daher vor dem Aufruf der Prozedur im Datensatz enthalten sein. Ausgegeben werden neben einem Protokoll des Minimierungsverlaufs die Schätzungen der Regressionsgewichte, Standardfehler, T-Werte und Signifikanzen der Parameter sowie die Effektkoeffizienten $\exp(\beta_k)$. Auf Anforderung werden auch die Korrelationen der Parameterschätzer ausgedruckt. Zur Beurteilung der Gesamterklärungskraft werden Pseudo-R^2-Werte (P^2) ausgegeben. Berechnet werden sowohl die relative Devianzreduktion gegenüber einem Nullmodell, das keine Regressionskonstanten enthält, als auch die relative Devianzreduktion gegenüber dem Konstantenmodell. Beim Nullmodell ohne Konstanten wird die Verbesserung des geschätzten logistischen Modells gegenüber einem Gleichverteilungsmodell gemessen, beim Konstantenmodell gegenüber einem Modell, das jedem Befragten die relativen Häufigkeiten der Randverteilung der abhängigen Variable als Wahrscheinlichkeiten zuordnet. Für beide Werte werden Likelihood-Ratio-Tests berechnet Zur Beurteilung der Prognosekraft des Modells in unterschiedlichen Kategorien wird sowohl eine Klassifikationstabelle als auch eine Vorhersagetabelle nach *McFadden* (1974) ausgedruckt. Schließlich ist es durch die Angabe von Kontrasten möglich, Wald-Tests berechnen zu lassen.

Die zu analysierenden Daten müssen in einer SPSS^x-Systemdatei enthalten sein. Im Unterschied zu anderen Programmen zur Schätzung logistischer Entscheidungsmodelle muß nicht für jede Untersuchungseinheit und jede Kategorie der abhängigen Variable eine eigene Datenzeile in der Datei definiert sein. Stattdessen geht die Prozedur von der üblichen quadratischen Datenmatrix aus, in der jede Zeile alle Informationen eines Falls enthält.

Der Aufruf des Macros erfolgt wie eine gewöhnliche SPSS^x-Prozedur. Allerdings sind die Anforderungen an die Einhaltung der Syntax strenger. Bei Fehlern gibt SPSS^X Fehlermeldungen der Prozedur MATRIX oder des Macro-Interpreters aus. Diese Fehlermeldungen sind in der Regel nicht zu deuten. Eigene Fehlermeldungen werden von dem SPSS^x-Macro ausgedruckt, wenn die Spezifikation von Wald-Tests inkonsistent ist sowie bei Schätzproblemen während der Minimierung der negativen logarithmierten Likelihood-Funktion.

Angefordert wird die Schätzung eines Modells durch den Aufruf:[1]

```
UTIL  VAR=variablenliste/ NKAT=kategorienzahl/ DEP=variablennummer /
      GENVAR=variablennummer, variablennummer, ... /
      ALTVAR=variablennummer, variablennummer, ... /
      ALTKAT=kategoriennummer, kategoriennummer, ... /
      COR /
      TEST=parameterzahl, parameterzahl, ... /
      DESIGN=wert, wert, ... /
```

[1] Die großgeschriebenen Wörter sind Schlüsselwörter, die unverändert übernommen werden müssen. Kleingeschriebene Wörter beziehen sich auf Parameterwerte, die je nach Anforderung unterschiedlich ausfallen.

NIT=iterationenzahl / EPS1=konvergenzwert1 / EPS2=konvergenzwert2 / NCASE=fallzahl /

Es ist nicht notwendig, alle Parameter des Aufrufs anzugeben. Notwendig ist allein die Angabe einer Liste von Variablen nach dem Schlüsselwort "VAR", die Angabe der Kategorienzahl der abhängigen Variable nach dem Schlüsselwort "NKAT" sowie die Angabe der Variablennummern generischer Variablen nach dem Schlüsselwort "GENVAR" und/oder die Angabe von Variablennummern für alternativenspezifische Variablen nach dem Schlüsselwort "ALTVAR". Werden alternativenspezifische Variablen spezifiziert, so sind nach dem Schlüsselwort "ALTKAT" die Angaben der Kategorien der abhängigen Variable vorzunehmen, auf die sich die alternativenspezifischen Variablen beziehen. Generell gilt die allgemeine Syntaxregel, daß die Angabe eines Parameters durch einen Schrägstrich (/) abgeschlossen wird.

Die Variablenliste nennt die Variablen der aktuellen Systemdatei *(active file)*, die in dem logistischen Modell berücksichtigt werden. Die "TO"-Konvention zur Abkürzung der Liste kann verwendet werden. Es ist wichtig, daß ein Variablenname in der Liste eindeutig ist. Wird zweimal derselbe Variablenname angegeben, bricht das Programm mit einer Fehlermeldung der $SPSS^X$-Prozedur MATRIX ab. Über die Variablenliste werden die Daten an das Programm übergeben. Ist ein Wert bei einer der Variablen der Variablenliste als ungültig *(MISSING)* gekennzeichnet, wird der gesamte Fall von der Analyse ausgeschlossen.

Die hinter dem Schlüsselwort "NKAT" anzugebende Anzahl der Kategorien der abhängigen Variable muß eine positive ganze Zahl größer 1 sein. Das Programm setzt voraus, daß die Ausprägungen der abhängigen Variable aus positiven ganzen Zahlen bestehen, deren kleinster Wert 1 ist. Es ist wichtig, daß auch nach dem Ausschluß fehlender Fälle alle Kategorien der abhängigen Variable besetzt sind. Anderenfalls ist das Programm nicht in der Lage, eine Lösung zu finden.

Standardmäßig wird angenommen, daß die erste Variable der Variablenliste die abhängige Variable ist. Ist dies nicht der Fall, kann hinter dem Schlüsselwort "DEP" die Variablennummer der abhängigen Variable angegeben werden. Variablennummern beziehen sich bei allen Parametern des Macros auf die Reihenfolge der Variablen in der angegebenen Variablenliste. Ist eine Variablennummer kleiner 1 oder größer der Anzahl der Variablen in der Liste, kommt es zu einem Fehlerabbruch.

Mit dem Schlüsselwort "GENKAT" kann eine Liste generischer Variablen festgelegt werden. Die hinter dem Schlüsselwort anzugebenen Zahlen bezeichnen wieder Variablennummern der Variablenliste. Da generische Variablen für jede Kategorie der abhängigen Variable eigene Ausprägungen haben, stehen hinter jeder generischen Variable des Modells I Variablen der Variablenliste, wobei I für die Zahl der Kategorien der abhängigen Variable steht. Es ist daher wichtig, daß die Anzahl der Variablennummern hinter dem Schlüsselwort "GENKAT" eine Zahl ist, die sich ohne Rest durch die Anzahl der Kategorien der abhängigen Variable teilen läßt. Das Macro geht dabei davon aus, daß sich die ersten I Variablennummern auf die erste generische Variable beziehen, die zweiten I Nummern auf die zweite generische Variable und so fort. Die einzelnen Variablennummern müssen durch jeweils genau ein Komma getrennt sein. Hinter der letzten Nummer darf kein Komma stehen.

Alternativenspezifische Variablen werden nach dem Schlüsselwort "ALTVAR" angegeben. Wie bei generischen Variablen sind auch hier Variablennummern durch genau ein Komma zu trennen. Neben den Variablennummern muß bei alternativenspezifischen Variablen auch angegeben werden, auf welche Kategorie der abhängigen Variable eine Variable wirkt. Dazu ist hinter dem Schlüsselwort "ALTKAT" eine Liste von Kategoriennummern anzugeben. Die einzelnen Kategoriennummern müssen wieder durch jeweils genau ein Komma getrennt sein. Wichtig ist, daß die Zahl der Kategoriennummern mit der Anzahl der alternativenspezifischen Variablen übereinstimmt.

Mit dem Schlüsselwort "COR /" können über die Standardausgabe hinaus Korrelationen der Schätzer angefordert werden. Zwischen dem Schlüsselwort und dem den Parameter abschließenden Schrägstrich muß mindestens ein Leerzeichen stehen.

Die Schlüsselwörter "TEST" und "DESIGN" fordern Wald-Tests an. Hinter dem Schlüsselwort "TEST" folgt eine Liste von Werten, die die Anzahl der Kontraste (Linearkombinationen der Regressionskoeffizienten) pro Test angibt. Die Werte müssen positive ganze Zahlen sein und dürfen nicht größer sein als die Anzahl der Parameter des geschätzten Modells. Jeder Wert steht für einen eigenen Test. Wird mehr als ein Test angefordert, so sind die Werte wieder jeweils durch ein Komma zu trennen. Hinter dem Schlüsselwort "DESIGN" sind die Designmatrizen der Tests zu spezifizieren. Dabei wird für jeden Kontrast zunächst das Gewicht angegeben, mit dem ein Regressionskoeffizient in den Kontrast einfließt. Die erste Zahl steht für das Gewicht des ersten Regressionskoeffizienten, die zweite Zahl für den zweiten Koeffizienten und so fort, bis allen Modellparametern des spezifizierten Modells ein Gewicht zugeordnet ist. Die Reihenfolge der Modellparameter ist dabei so, daß zuerst das Regressionsgewicht der ersten generischen Variable kommt, dann das Gewicht der zweiten generischen Variable und so fort. Nach den generischen Variablen folgen die Gewichte der alternativenspezifischen Variablen in der Reihenfolge, wie sie hinter dem Schlüsselwort "ALTVAR" aufgelistet sind. Nach der Angabe der Gewichte eines Kontrastes folgt der Wert, den der zu testende Kontrast nach der Nullhypothese in der Grundgesamtheit haben soll. Nach dem ersten Kontrast folgen die weiteren Kontraste eines Tests und anschließend eventuell die von weiteren Tests. Die einzelnen Zahlen der Werteliste nach dem Schlüsselwort "DESIGN" sind wieder durch Kommata zu trennen.

Mit den Parametern hinter den optionalen Schlüsselwörtern "NIT", "EPS1" und "EPS2" kann der Schätzalgorithmus beeinflußt werden. Die Zahl hinter dem Schlüsselwort "NIT" gibt die maximale Iterationszahl an. Ist die Lösung bis zu dieser Iteration noch nicht konvergiert, wird eine entsprechende Meldung ausgegeben. Voreingestellt ist als maximale Iterationszahl 25. Das Erreichen der Konvergenz wird durch die beiden Abbruchkriterien EPS1 und EPS2 gesteuert. Konvergenz wird angenommen, wenn die relative Devianzreduktion zwischen zwei Iterationen kleiner oder gleich EPS 1 ist oder das betragsmäßig größte Element des Vektors der ersten Ableitungen der Minimierungsfunktion nach den Modellparametern kleiner oder gleich EPS2 ist. Voreingestellt ist für EPS1 der Wert 10^{-7} und für EPS2 der Wert 10^{-4}.

Vor Beginn der Minimierung arbeitet das Programm in einer Schleife alle Fälle genau einmal einzeln ab. Enthält die Stichprobe mehr als 20 000 Fälle, so ist dies dem Macro durch eine entsprechend hohe Zahl nach dem Schlüsselwort "NCASE" mitzuteilen.

Zur Verdeutlichung der Spezifikation eines Modells will ich die Spezifikation des in Abbildung 8 wiedergegebenen Modells vorstellen. Die Variablennamen seien "BEHAV" für die abhängige

Variable, "INTKOOP", "INTAUSL", "INTLUEG" und "INTBOYK" für die erklärenden Dummy-Variablen und "EINS" für eine Konstante mit dem Wert 1. Bei der abhängigen Variable steht der Wert 1 für "Kooperation", 2 für "Unvollständigkeit", 3 für "falsche Angaben" und 4 für "Boykott". Neben der Schätzung der Modellparameter soll ein Wald-Test die Gleichheit der vier alternativenspezifischen Regressionskoeffizienten testen. Zusätzlich zur Standardausgabe sollen auch die Korrelationen der Parameterschätzungen ausgegeben werden. Der Aufruf des Macros sieht dann folgendermaßen aus:

```
UTIL  VAR=BEHAV,INTKOOP,INTAUSL,INTLUEG,INTBOYK,EINS
      /NKAT=4 /DEP=1
      /ALTVAR=6,6,6,2,3,4,5
      /ALTKAT=1,2,3,1,2,3,4
      /COR
      /TEST=3
      /DESIGN=     0,0,0,1,0,0,-1,0,
                   0,0,0,0,1,0,-1,0,
                   0,0,0,0,0,1,-1,0/
```

Nach der Angabe der im Modell verwendeten Parameter wird dem Programm durch die Angabe von "NKAT=4" mitgeteilt, daß die abhängige Variable vier Kategorien hat. Als abhängige Variable wird die erste Variable der Variablenliste identifiziert ("DEP=1"). Hinter dem Schlüsselwort "ALTVAR=" werden 7 alternativenspezifische Variablen spezifiziert. Die ersten drei Variablen sprechen die 6. Variable der Variablenliste, also die Konstante "EINS", an. Sie definieren die alternativenspezifischen Regressionskonstanten. Anschließend werden die vier 0/1-kodierten Dummy-Variablen als alternativenspezifische Variablen definiert. Hinter dem Schlüsselwort "ALTKAT=" sind die Kategorien aufgeführt, die die alternativenspezifischen Variablen beeinflussen. Die drei alternativenspezifischen Regressionskonstanten betreffen die ersten drei Kategorien von "BEHAV". Die nachfolgenden Dummy-Variablen der Handlungsabsicht wirken auf die jeweils korrespondierende Kategorie des berichteten Verhaltens. Über diese Spezifikation des Modells ist auch die Reihenfolge der Regressionskoeffizienten festgelegt. Danach sind die ersten drei Modellparameter alternativenspezifische Konstanten für die ersten drei Kategorien der abhängigen Variable und die nachfolgenden Parameter die Regressionsgewichte der vier Dummy-Variablen. Da das Schlüsselwort "COR" im Aufruf vorkommt, werden auch die Korrelationen der Schätzungen berechnet und ausgegeben.

Die Zahl "3" hinter dem Schlüsselwort "TEST" gibt an, daß ein Wald-Test durchgeführt werden soll, bei dem drei Kontraste simultan getestet werden sollen. Die Gewichte und die in der Grundgesamtheit erwarteten Werte sind nach dem Schlüsselwort "DESIGN" angegeben. Die ersten 7 Werte geben die Gewichte des ersten Kontrastes an. Die Werte besagen, daß der vierte Modellparameter mit einem positiven Gewicht von +1 und der siebte Modellparameter mit einem Gewicht von -1 in den Kontrast eingehen. Dieser Kontrast ist also die Differenz $\beta_{Koop.}-\beta_{Boyk.}$ Der achte Wert in der ersten Zeile nach dem Schlüsselwort "DESIGN" postuliert, daß diese Differenz Null sein soll. Analog werden in den folgenden zwei Zeilen die Kontraste $\beta_{Unv.}-\beta_{Boyk.}=0$ und $\beta_{falsch.}-\beta_{Boyk.}=0$ spezifiziert. Es wird also folgende Nullhypothese getestet:

$$H_0: \beta_{Koop.}-\beta_{Boyk.}=0 \;\&\; \beta_{Unv.}-\beta_{Boyk.}=0 \;\&\; \beta_{falsch.}-\beta_{Boyk.}=0.$$

Die Hypothese postuliert, daß alle Regressionsgewichte der vier alternativenspezifischen Variablen den gleichen Wert haben.