

Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr

Inauguraldissertation

zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften
durch die Wirtschaftswissenschaftliche Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von

Cornelia Richter
aus Essen

Erster Berichterstatter: Prof. Dr. Karl-Hans Hartwig
Zweiter Berichterstatter: Prof. Dr. Dr. h.c. Klaus Backhaus
Dekan: Prof. Dr. Christoph Watrin
Tag der mündlichen Prüfung: 14. November 2013

Cornelia Richter

**Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl
im Regionalverkehr**

Wissenschaftliche Schriften der WWU Münster

Reihe IV

Band 7

Cornelia Richter

Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr

Wissenschaftliche Schriften der WWU Münster

herausgegeben von der Universitäts- und Landesbibliothek Münster
<http://www.ulb.uni-muenster.de>

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Dieses Buch steht gleichzeitig in einer elektronischen Version über den Publikations- und
Archivierungsserver der WWU Münster zur Verfügung.
<http://www.ulb.uni-muenster.de/wissenschaftliche-schriften>

Cornelia Richter

„Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr“
Wissenschaftliche Schriften der WWU Münster, Reihe IV, Band 7

© 2014 der vorliegenden Ausgabe:

Die Reihe „Wissenschaftliche Schriften der WWU Münster“ erscheint im Verlagshaus Monsenstein und
Vannerdat OHG Münster
www.mv-wissenschaft.com

ISBN 978-3-8405-0094-7 (Druckausgabe)

URN urn:nbn:de:hbz:6-44319583036 (elektronische Version)

direkt zur Online-Version:

© 2014 Cornelia Richter
Alle Rechte vorbehalten

Satz: Cornelia Richter
Titelbild: Cornelia Richter
Umschlag: MV-Verlag
Druck und Bindung: MV-Verlag



Danksagung

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit als wissenschaftliche Mitarbeiterin im Projekt „Qualitätssimulationsmodell im Regionalverkehr“ am Forschungsschwerpunkt Stadtverkehr der Westfälischen Hochschule. Sie wurde im September 2013 als Dissertation an der Westfälischen Wilhelms-Universität Münster angenommen.

Ich danke an dieser Stelle allen, die mich in meiner Promotionszeit in vielfältiger Art und Weise unterstützt haben. Mein Dank gilt insbesondere

- meinem Doktorvater, Herrn Prof. Dr. Karl-Hans Hartwig, für die Betreuung und Begutachtung dieser Arbeit sowie das große Vertrauen, das er mir entgegengebracht hat,
- Herrn Prof. Dr. Dr. h.c. Klaus Backhaus für die freundliche Übernahme des Zweitgutachtens,
- Herrn Prof. Dr. Stephan Keuchel für die vielen anregenden Diskussionen, die hilfreiche Unterstützung sowie die großen Freiheiten, die sehr zum Gelingen dieser Arbeit beigetragen haben,
- den Mitarbeitern der eurobahn, der NordWestBahn und der WestfalenBahn, insbesondere Frau Doris Reich und den Herren Christian Arndt, Rainer Blüm, Thomas Kopp und Henning Nüßgen für die nützlichen Ratschläge aus der Praxis und die Möglichkeit zur Durchführung von Fahrgastbefragungen in den Zügen,
- den studentischen und wissenschaftlichen Hilfskräften für das große Engagement bei den Befragungen,
- den Teilnehmern der Konferenzen und Workshops, insbesondere Jordan Louviere, Eric Molin, Harmen Oppewal und Tiago Ribeiro für den wissenschaftlichen Austausch und die hilfreichen Ratschläge auch außerhalb der publizierten Literatur,
- dem BMBF für die Förderung von „QUASIMIR“, wodurch die große empirische Datengrundlage und die Teilnahme an den Konferenzen möglich geworden ist,
- meinen Kollegen vom Fachbereich Wirtschaftsingenieurwesen für die Hilfsbereitschaft und das gute Arbeitsumfeld im Hochschulalltag,
- meinem Freundeskreis für das Leben außerhalb von Diss und Arbeit

und meiner Familie für die große Unterstützung, insbesondere meinen Eltern, meinem Bruder und natürlich ganz besonders William.

Frankfurt am Main, im Frühjahr 2014

Conny Richter

Inhaltsverzeichnis

| | |
|--|-----------|
| Tabellenverzeichnis | V |
| Abbildungsverzeichnis | IX |
| Abkürzungsverzeichnis | XI |
| 1 Einleitung..... | 1 |
| 1.1 Problemstellung..... | 1 |
| 1.2 Zielsetzung und Aufbau der Arbeit..... | 3 |
| 2 Qualität im Schienenpersonennahverkehr | 7 |
| 2.1 Rahmenbedingungen im Schienenpersonennahverkehr | 7 |
| 2.1.1 Rahmenbedingungen auf europäischer Ebene | 7 |
| 2.1.2 Rahmenbedingungen auf nationaler Ebene | 8 |
| 2.1.3 Rahmenbedingungen auf regionaler Ebene..... | 10 |
| 2.1.4 Nahverkehrsplan des Zweckverbands Nahverkehr Westfalen-Lippe | 11 |
| 2.1.5 Vertragsformen und Anreizstrukturen..... | 15 |
| 2.2 Qualität und Kundenzufriedenheit | 18 |
| 2.2.1 DIN EN 13816..... | 18 |
| 2.2.2 Verfahren zur Messung der Qualität und der Kundenzufriedenheit..... | 20 |
| 2.2.3 Kundenzufriedenheit und wirtschaftlicher Erfolg..... | 26 |
| 2.2.4 Bildung eines Qualitätsindexes | 30 |
| 2.2.5 Kundenzufriedenheit, Qualität und Verkehrsmittelwahl | 32 |
| 2.3 Zusammenfassung..... | 33 |
| 3 Methodische Grundlagen der Discrete Choice Analyse | 37 |
| 3.1 Allgemeine Grundlagen der Discrete Choice Analyse | 37 |
| 3.2 Binomiales und multinomiales Logit-Modell..... | 41 |
| 3.3 Nested Logit-Modell | 48 |
| 3.4 Abgrenzung zu komplexeren Modellen der Discrete Choice Analyse | 53 |

| | | |
|----------|--|-----------|
| 3.5 | Testverfahren..... | 54 |
| 3.5.1 | Hausman-McFadden-Test der Unabhängigkeit von irrelevanten Alternativen..... | 54 |
| 3.5.2 | Likelihood-Ratio-Test..... | 55 |
| 3.5.3 | Modellgüte und Validität..... | 62 |
| 3.6 | Zusammenfassung..... | 66 |
| 4 | Verfahren zur Erhebung bei komplexen Multi-Attribut-Entscheidungen..... | 69 |
| 4.1 | Allgemeine Verfahren..... | 70 |
| 4.2 | Verfahren der hierarchischen Informationsintegration (HII)..... | 72 |
| 4.2.1 | Theorie der Informationsintegration..... | 72 |
| 4.2.2 | Grundannahmen der HII..... | 73 |
| 4.2.3 | Aufbau der Experimente der HII..... | 75 |
| 4.2.4 | Analyse der Experimente der HII..... | 77 |
| 4.2.5 | Verfahrensvarianten der HII..... | 79 |
| 4.2.6 | Testverfahren..... | 80 |
| 4.2.7 | Kritik an dem Verfahren der HII..... | 83 |
| 4.3 | Verfahren der integrierten hierarchischen Informationsintegration (HII-I)..... | 85 |
| 4.3.1 | Aufbau der Experimente der HII-I..... | 85 |
| 4.3.2 | Analyse der Experimente der HII-I..... | 86 |
| 4.3.3 | Testverfahren..... | 88 |
| 4.4 | Anwendungsbeispiele der Verfahren der HII und der HII-I..... | 90 |
| 4.5 | Zusammenfassung..... | 97 |
| 5 | Verfahren zur empirischen Entwicklung von Konstrukten..... | 99 |
| 5.1 | Verfahren zur Erhebung von Ähnlichkeitsdaten..... | 99 |
| 5.2 | Multidimensionale Skalierung..... | 104 |
| 5.3 | Vergleich von Partitionen..... | 113 |
| 5.4 | Anwendungsbeispiel zur empirischen Entwicklung von Konstrukten..... | 115 |
| 5.5 | Zusammenfassung..... | 116 |

| | | |
|----------|---|------------|
| 6 | Empirische Entwicklung von Konstrukten für die Anwendung des Verfahrens der HII-I im Regionalverkehr | 119 |
| 6.1 | Untersuchungsdesign und Datenerhebung..... | 121 |
| 6.2 | Beschreibung der Stichprobe | 126 |
| 6.3 | Analyse der Unähnlichkeitsdaten..... | 128 |
| 6.4 | Interpretation und Überprüfung der Konfiguration | 131 |
| 6.5 | Analyse der Ratingdaten | 137 |
| 6.6 | Auswahl von Konstrukten und Attributen | 141 |
| 6.7 | Zusammenfassung..... | 142 |
| 7 | Anwendung des Verfahrens der HII-I im Regionalverkehr | 145 |
| 7.1 | Untersuchungsdesign und Datenerhebung..... | 145 |
| 7.2 | Beschreibung der Stichprobe | 158 |
| 7.3 | Analyse der Rating-Experimente | 165 |
| 7.4 | Aufbau der Analyse der hierarchischen Wahlexperimente | 170 |
| 7.5 | Multinomiales Logit-Modell..... | 172 |
| 7.5.1 | Subexperiment der Verbindungsqualität | 173 |
| 7.5.2 | Subexperiment des Komforts | 181 |
| 7.5.3 | Subexperiment der Informationen | 186 |
| 7.6 | Überprüfung der Gleichheit der Nutzenfunktionen der Nutzergruppen I..... | 190 |
| 7.6.1 | Nicht wahlfreie Bahnnutzer und wahlfreie Bahnnutzer..... | 191 |
| 7.6.2 | Wahlfreie Bahnnutzer und „wahlfreie“ Pkw-Nutzer | 192 |
| 7.7 | Hausman-McFadden-Test der Unabhängigkeit von irrelevanten Alternativen..... | 194 |
| 7.7.1 | Wahlfreie Bahnnutzer | 194 |
| 7.7.2 | „Wahlfreie“ Pkw-Nutzer | 195 |
| 7.8 | Nested Logit-Modell | 195 |
| 7.8.1 | Subexperiment der Verbindungsqualität | 196 |
| 7.8.2 | Subexperiment des Komforts | 199 |
| 7.8.3 | Subexperiment der Informationen | 201 |

| | | |
|---------------|--|------------|
| 7.9 | Überprüfung der Gleichheit der Nutzenfunktionen der Nutzergruppen II | 203 |
| 7.9.1 | Nicht wahlfreie und wahlfreie Bahnnutzer | 203 |
| 7.9.2 | Wahlfreie Bahnnutzer und „wahlfreie“ Pkw-Nutzer | 206 |
| 7.10 | Überprüfung der Prozessgleichheit zwischen den Subexperimenten | 207 |
| 7.10.1 | Nicht wahlfreie Bahnnutzer | 208 |
| 7.10.2 | Wahlfreie Bahnnutzer | 212 |
| 7.10.3 | „Wahlfreie“ Pkw-Nutzer | 216 |
| 7.11 | Gemeinsames Modell | 219 |
| 7.11.1 | Nicht wahlfreie Bahnnutzer | 219 |
| 7.11.2 | Wahlfreie Bahnnutzer | 220 |
| 7.11.3 | „Wahlfreie“ Pkw-Nutzer | 221 |
| 7.12 | Bestimmung von Zahlungsbereitschaften | 223 |
| 7.13 | Simulation von Wahlwahrscheinlichkeiten | 228 |
| 7.14 | Zusammenfassung | 238 |
| 8 | Fazit und Ausblick..... | 245 |
| Anhang | | 249 |
| A. | Skalenniveau der Daten der aggregierten Unähnlichkeitsmatrix | 249 |
| B. | Daten der aggregierten Unähnlichkeitsmatrix | 251 |
| C. | Erhebungsinstrument zur empirischen Entwicklung der Konstrukte | 253 |
| D. | Erhebungsinstrument mit HII-I für Bahnnutzer | 257 |
| E. | Erhebungsinstrument mit HII-I für Pkw-Nutzer | 265 |
| | Literaturverzeichnis..... | 271 |
| | Quellenverzeichnis | 287 |

Tabellenverzeichnis

| | | |
|------------|--|-----|
| Tab. 4.1: | Anwendungen des Verfahrens der HII und der HII-I..... | 94 |
| Tab. 5.1: | Objektpaare zum Vergleich von Partitionen | 113 |
| Tab. 6.1: | Qualitätsmerkmale zur empirischen Entwicklung von Konstrukten..... | 121 |
| Tab. 6.2: | Absolute und relative adjustierte Häufigkeiten der Charakteristika zur Soziodemografie und zum allgemei- nen Verkehrsverhalten differenziert nach Nutzergruppen .. | 127 |
| Tab. 6.3: | Anpassungsmaße der ermittelten Konfigurationen | 130 |
| Tab. 6.4: | Cluster und Qualitätsmerkmale | 133 |
| Tab. 6.5: | Lage- und Streuungsmaße der Ratingdaten differenziert nach Bahnnutzern und Pkw-Nutzern..... | 138 |
| Tab. 7.1: | Ausprägungen der qualitativen Attribute der Konstrukte Verbindungsqualität, Komfort und Informationen..... | 147 |
| Tab. 7.2: | Absolute und relative adjustierte Häufigkeiten der Charakteristika zur Soziodemografie und zum allgemeinen Verkehrsverhalten differenziert nach Nutzergruppen..... | 159 |
| Tab. 7.3: | Gesamtreisezeiten und Fahrtkosten der aktuellen bzw. alternativen Fahrten differenziert nach Nutzergruppen | 161 |
| Tab. 7.4: | Effektkodierung des Attributs Pünktlichkeit | 166 |
| Tab. 7.5: | Regressionsmodell der Verbindungsqualität differenziert nach Nutzergruppen..... | 167 |
| Tab. 7.6: | Regressionsmodell des Komforts differenziert nach Nutzergruppen | 168 |
| Tab. 7.7: | Regressionsmodell der Informationen differenziert nach Nutzergruppen | 169 |
| Tab. 7.8: | Multinomiales Logit-Modell der Verbindungsqualität für die nicht wahlfreien und die wahlfreien Bahnnutzer | 174 |
| Tab. 7.9: | Multinomiales Logit-Modell der Verbindungsqualität für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer | 175 |
| Tab. 7.10: | Überprüfung der Gesamtsignifikanz des MNL-Modells der Verbindungsqualität differenziert nach Nutzergruppen | 176 |

| | |
|---|-----|
| Tab. 7.11: Multinomiales Logit-Modell des Komforts für die nicht wahlfreien und die wahlfreien Bahnnutzer..... | 182 |
| Tab. 7.12: Multinomiales Logit-Modell des Komforts für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer..... | 183 |
| Tab. 7.13: Überprüfung der Gesamtsignifikanz des MNL-Modells des Komforts differenziert nach Nutzergruppen..... | 183 |
| Tab. 7.14: Multinomiales Logit-Modell der Informationen für die nicht wahlfreien und die wahlfreien Bahnnutzer..... | 187 |
| Tab. 7.15: Multinomiales Logit-Modell der Informationen für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer..... | 188 |
| Tab. 7.16: Überprüfung der Gesamtsignifikanz des MNL-Modells der Informationen differenziert nach Nutzergruppen..... | 188 |
| Tab. 7.17: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten..... | 192 |
| Tab. 7.18: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer differenziert nach Subexperimenten..... | 193 |
| Tab. 7.19: Nested Logit-Modell der Verbindungsqualität für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer | 197 |
| Tab. 7.20: Nested Logit-Modell des Komforts für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer..... | 199 |
| Tab. 7.21: Nested Logit-Modell der Informationen für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer | 201 |
| Tab. 7.22: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten..... | 205 |
| Tab. 7.23: Zweite Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten..... | 206 |

| | |
|---|-----|
| Tab. 7.24: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer differenziert nach Subexperimenten..... | 207 |
| Tab. 7.25: Gemeinsame Modelle mit allen Subexperimenten der nicht wahlfreien Bahnnutzer..... | 210 |
| Tab. 7.26: Gemeinsame Modelle mit allen Subexperimenten der wahlfreien Bahnnutzer..... | 214 |
| Tab. 7.27: Gemeinsame Modelle mit allen Subexperimenten der „wahlfreien“ Pkw-Nutzer..... | 217 |
| Tab. 7.28: Marginale Zahlungsbereitschaften differenziert nach Nutzergruppen..... | 224 |
| Tab. 7.29: Spannweiten der marginalen Zahlungsbereitschaften differenziert nach Nutzergruppen..... | 226 |
| Tab. 7.30: Szenarien der Simulation für die nicht wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer..... | 229 |
| Tab. 7.31: Mittelwert der Wahlwahrscheinlichkeiten der Alternativen differenziert nach den nicht wahlfreien Bahnnutzern und den „wahlfreien“ Pkw-Nutzern..... | 231 |
| Tab. 7.32: Szenarien der Simulation für die wahlfreien Bahnnutzer ... | 234 |
| Tab. 7.33: Mittelwert der Wahlwahrscheinlichkeiten der Alternativen der wahlfreien Bahnnutzer..... | 237 |

Abbildungsverzeichnis

| | | |
|-----------|--|-----|
| Abb. 2.1: | Qualitätskreis | 19 |
| Abb. 2.2: | Systematisierung nachfragerbezogener Verfahren zur Messung der Qualität bzw. der Zufriedenheit | 22 |
| Abb. 3.1: | Grundsätzlicher Aufbau eines Nested Logit-Modells mit vier Ebenen | 48 |
| Abb. 3.2: | Exemplarische Struktur der unrestringierten Modelle und des restringierten Modells zur Überprüfung der ersten Subhypothese des zweistufigen Likelihood-Ratio-Tests | 60 |
| Abb. 3.3: | Exemplarische Struktur des unrestringierten und des restringierten Modells zur Überprüfung der zweiten Subhypothese des zweistufigen Likelihood-Ratio-Tests | 61 |
| Abb. 4.1: | Grundsätzlicher Aufbau eines Experiments der HII | 76 |
| Abb. 4.2: | Grundsätzlicher Aufbau eines Experiments der HII-I..... | 86 |
| Abb. 6.1: | Scree-Plot mit normalisiertem Rohstress | 130 |
| Abb. 6.2: | Orthogonale Projektion der ausgewählten Konfiguration x-y-Ebene..... | 131 |
| Abb. 6.3: | Orthogonale Projektion der ausgewählten Konfiguration x-z-Ebene | 132 |
| Abb. 6.4: | Orthogonale Projektion der ausgewählten Konfiguration z-y-Ebene | 132 |
| Abb. 6.5: | Verteilung der Rand-Indizes – Vergleich der Sorting- Daten der Bahnnutzer mit der aus fünf Clustern und fünf einzelnen Merkmalen bestehenden Partition..... | 135 |
| Abb. 6.6: | Verteilung der Rand-Indizes – Vergleich der Sorting- Daten der Bahnnutzer mit der aus sieben Clustern und fünf einzelnen Merkmalen bestehenden Partition | 135 |
| Abb. 6.7: | Verteilung der Rand-Indizes – Vergleich der Sorting- Daten der Pkw-Nutzer mit der aus fünf Clustern und fünf einzelnen Merkmalen bestehenden Partition..... | 137 |
| Abb. 7.1: | Aufbau des eingesetzten Erhebungsinstruments zur Befragung der Nutzer des regionalen SPNV | 149 |
| Abb. 7.2: | Aufbau des Subexperiments der integrierten hierarchischen Informationsintegration..... | 152 |

| | | |
|------------|--|-----|
| Abb. 7.3: | Aufbau des eingesetzten Erhebungsinstruments zur Befragung der Nicht-Nutzer des regionalen SPNV (Befragung in den Behörden) | 155 |
| Abb. 7.4: | Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt der Verbindungsqualität differenziert nach Nutzergruppen | 162 |
| Abb. 7.5: | Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt des Komforts differenziert nach Nutzergruppen | 163 |
| Abb. 7.6: | Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt der Informationen differenziert nach Nutzergruppen | 164 |
| Abb. 7.7: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer | 191 |
| Abb. 7.8: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer | 193 |
| Abb. 7.9: | Aufbau der Nested Logit-Modelle | 196 |
| Abb. 7.10: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer | 204 |
| Abb. 7.11: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer | 206 |
| Abb. 7.12: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien Bahnnutzer zwischen den Subexperimenten | 209 |
| Abb. 7.13: | Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer zwischen den Subexperimenten | 212 |

Abkürzungsverzeichnis

| | |
|---------------|---|
| AEG | Allgemeines Eisenbahngesetz |
| DIN | Deutsche Industrie-Norm |
| EG | Europäische Gemeinschaft |
| EN | Europäische Norm |
| EWG | Europäische Wirtschaftsgemeinschaft |
| FP-Experiment | „Full-Profile“-Experiment |
| GEV | generalisierter Extremwert (<i>generalised extreme value</i>) |
| HII | hierarchische Informationsintegration |
| HII-I | integrierte hierarchische Informationsintegration |
| IIA | Unabhängigkeit von irrelevanten Alternativen (<i>independence of/from irrelevant alternatives</i>) |
| IID | unabhängig und identisch verteilt (<i>independently and identically distributed</i>) |
| IQR | Interquartilsabstand |
| IV | Individualverkehr |
| IV-Parameter | Inklusivwertparameter |
| MDS | multidimensionale Skalierung |
| MiD | Mobilität in Deutschland |
| MIV | mobilisierter Individualverkehr |
| MNL-Modell | multinomiales Logit-Modell |
| NE | Nutzeneinheit/en |
| NL-Modell | Nested Logit-Modell |
| NWL | Nahverkehr Westfalen-Lippe (Zweckverband) |
| ÖPNV | öffentlicher Personennahverkehr |
| ÖPV | öffentlicher Personenverkehr |
| ÖV | öffentlicher Verkehr |
| RegG | Regionalisierungsgesetz |
| SD | Standardabweichung |
| SPNV | Schienenpersonennahverkehr |

1 Einleitung

1.1 Problemstellung

Mobilität wird allgemein als ein Grundbedürfnis von Menschen und als Voraussetzung für die Teilnahme am gesellschaftlichen Leben angesehen. Im Sinne der Daseinsvorsorge lässt sich die Bereitstellung eines Grundangebots des öffentlichen Personennahverkehrs (ÖPNV) durch die öffentliche Hand rechtfertigen, um ein Mindestmaß an Mobilität zu gewährleisten, da in einem vollkommen liberalisierten Markt Verkehrsdienstleistungen im Bereich des ÖPNV nicht oder nur zu sehr hohen Preisen angeboten würden.¹

Über dieses Grundangebot hinaus lässt sich eine lenkungspolitisch motivierte Bereitstellung eines qualitativ höherwertigen ÖPNV begründen, um durch eine Verkehrsverlagerung beispielsweise umweltpolitische Ziele zu verfolgen.² Dieses setzt allerdings voraus, dass die Qualität des ÖPNV tatsächlich die Verkehrsmittelwahl beeinflusst.

Die Bereitstellung durch die öffentliche Hand bedeutet nicht, dass die Verkehrsleistungen durch die öffentliche Hand selbst erbracht werden; sie können auch bei privaten oder öffentlichen Unternehmen bestellt werden. Sofern Ausgleichsleistungen und/oder ausschließliche Rechte für die Erbringung der Verkehrsdienstleistungen vereinbart werden, ist ein öffentlicher Dienstleistungsauftrag notwendig.³ Abhängig von der dabei verwendeten Vertragsform besteht für die Verkehrsunternehmen ein mehr

¹ Vgl. Schäfer-Sparenberg/Bongardt/Dalkmann (2006), S. 49, Holz-Rau et al. (2010), S. 5–6 und Schneiderhan (2012), S. 20.

² Vgl. Bundesministerium für Verkehr, Bau und Stadtentwicklung (2011), S. 23, Umweltbundesamt Dessau (2005), S. 42–43 und Forschungsgesellschaft für Straßen- und Verkehrswesen (2006), S. 6.

Aus ökonomischer Sicht wäre eine Internalisierung der externen Effekte des Verkehrs anzustreben, sodass bei der Verkehrsmittelwahl alle Kosten in das Entscheidungskalkül einbezogen würden. Da sich die einzelnen Verkehrsträger hinsichtlich der mit ihnen verbundenen externen Kosten unterscheiden, wäre durch eine Internalisierung eine wohlfahrtsmaximierende Veränderung des Modal Splits zu erwarten und einzelne lenkungspolitische Maßnahmen würden obsolet. Eine weitreichende Internalisierung ist derzeit jedoch nicht zu erwarten.

³ Vgl. § 19 Verordnung (EG) Nr. 1370/2007.

oder weniger stark ausgeprägter ökonomischer Anreiz zur Steigerung von Fahrgastzahlen.

Ist ein solcher Anreiz gegeben, werden Verkehrsunternehmen nur ökonomisch sinnvolle Maßnahmen zur Steigerung der Qualität durchführen, d. h. Maßnahmen, deren Kosten geringer sind als die zusätzlichen Einnahmen infolge gestiegener Fahrgastzahlen. Sofern ein solcher Anreiz nicht gegeben ist, jedoch ein qualitativ hochwertiger ÖPNV lenkungs- politisch gewünscht ist, muss das zu erbringende Qualitätsniveau entsprechend vertraglich vereinbart und dessen Erfüllung kontrolliert werden. Zu diesem Zweck werden in der Regel Qualitäts- bzw. Kundenzufriedenheitsmessungen durchgeführt.

Die Festlegung eines gewünschten Qualitätsniveaus sollte im Sinne eines effizienten Einsatzes öffentlicher Mittel erfolgen, sodass nur solche Maßnahmen zur Steigerung der Qualität angestrebt werden, die tatsächlich einen Einfluss auf die Verkehrsmittelwahl ausüben. Der Zusammenhang zwischen Qualität und Verkehrsmittelwahl ist somit je nach Vertragsform aus einzel- und/oder gemeinwirtschaftlicher Sicht von Bedeutung.

Bei gängigen Qualitäts- und Kundenzufriedenheitsmessinstrumenten wird zwar mitunter die Wichtigkeit der einzelnen Qualitätsmerkmale erfragt oder abgeleitet, allerdings erfolgt keine Abwägung zwischen den einzelnen Merkmalen, sodass im Extremfall eine hohe Qualität bei einem geringen Fahrpreis gefordert wird. Zudem ist ein direkter Bezug zur Verkehrsmittelwahl nicht gegeben.⁴

Bei der Verwendung eines Verfahrens der Discrete Choice Analyse kann ein direkter Bezug zur Verkehrsmittelwahl hergestellt werden. Dabei lässt sich ausgehend von Wahlentscheidungen der Einfluss von einzelnen Attributen⁵ bzw. deren Ausprägungen auf die Verkehrsmittelwahl bestimmen. Grundsätzlich könnte eine solche Analyse auf Realdaten basieren. Dabei besteht jedoch eine Beschränkung auf die in der Realität vorzufindenden Verkehrsmittel und deren Qualität, sodass keine differenzierte Aussage

⁴ Vgl. *Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006), S. 8–9.

⁵ Im Rahmen dieser Arbeit werden (Qualitäts-)Merkmale, die in Verkehrsmittelwahlmodellen verwendet werden, als (Qualitäts-)Attribute bezeichnet.

über den Zusammenhang zwischen Qualität und Verkehrsmittelwahl möglich ist.

Diese Beschränkung besteht bei Verkehrsmittelwahlexperimenten nicht. In fiktiven Wahlsituationen wählen Befragte zwischen Verkehrsmitteln, die durch Attribute mit unterschiedlichen Ausprägungen charakterisiert werden. Allerdings ist die Anzahl der Attribute (und deren Ausprägungen) zur Charakterisierung der Alternativen bei der Datenerhebung mit gängigen Wahlexperimenten aufgrund der Komplexität der Experimente begrenzt.⁶

In Discrete Choice Experimenten, die auf dem Verfahren der hierarchischen Informationsintegration (HII) und dessen Weiterentwicklung, der integrierten hierarchischen Informationsintegration (HII-I), beruhen, kann eine größere Anzahl von Attributen berücksichtigt werden.⁷ Dabei werden (ähnliche) Attribute zu Konstrukten zusammengefasst und in die hierarchischen Wahlexperimente einbezogen. Die dem Experiment zugrunde liegende hierarchische Struktur kann bei dem Verfahren der HII-I, nicht aber bei dem Verfahren der HII, überprüft werden.

Die Anzahl der Anwendungen des Verfahrens der HII-I, insbesondere aber der Anwendungen, bei denen die zugrunde liegende hierarchische Struktur überprüft wurde, ist sehr gering, sodass bislang die empirische Evidenz für dieses Verfahren fehlt.

1.2 Zielsetzung und Aufbau der Arbeit

Der vorliegenden Arbeit liegen zwei Ziele zugrunde. Das erste Ziel besteht darin, den Einfluss der Qualität des regionalen ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr zu analysieren. Zu diesem Zweck wird das Verfahren der integrierten hierarchischen Informationsintegration verwendet, um eine größere Anzahl von Attributen in Verkehrsmittelwahlexperimenten zu berücksichtigen.

⁶ Vgl. *Molin/Timmermans* (2009), S. 635, *Timmermans/Molin* (2009), S. 561 und *Pullman/Dodson/Moore* (1999), S. 126.

⁷ Vgl. *Louviere* (1984), S. 148–151, *Oppewal/Louviere/Timmermans* (1994), S. 93–95 und *Molin/Timmermans* (2009), S. 635–637.

Ein Schwerpunkt liegt dabei auf dem regionalen Schienenpersonen-nahverkehr (SPNV). Die Attribute zur Beschreibung der öffentlichen Verkehrsmittel werden so ausgewählt, dass sie weitestgehend durch die Verkehrsunternehmen beeinflussbar sind. Zur Analyse des Einflusses der Qualität des ÖPNV auf die Verkehrsmittelwahl werden in Befragungen insbesondere von Nutzern, aber auch von Nicht-Nutzern des regionalen SPNV in der Region Münster – Bielefeld – Osnabrück Verkehrsmittelwahl-experimente durchgeführt. Dabei stehen als Wahlalternativen eine Regionalbahn, ein Regionalbus und ggf. ein Pkw zur Auswahl.

Das zweite Ziel dieser Arbeit liegt darin, für diese Anwendung der integrierten hierarchischen Informationsintegration die hierarchische Struktur zu überprüfen, um einen Beitrag für die empirische Evidenz des Verfahrens zu leisten.

Die vorliegende Arbeit ist folgendermaßen aufgebaut: Zunächst werden in Kapitel 1 der institutionelle Rahmen des regionalen SPNV dargelegt sowie bestehende Qualitäts- und Kundenzufriedenheitsmessinstrumente vorgestellt. Darüber hinaus werden exemplarisch drei sehr unterschiedliche Studien mit Bezug zur Verkehrsmittelwahl bzw. zum wirtschaftlichen Erfolg beschrieben und kritisch gewürdigt. In Kapitel 3 erfolgt eine Beschreibung der methodischen Grundlagen der Discrete Choice Analyse. Darauf aufbauend werden in Kapitel 4 die Verfahren der hierarchischen Informationsintegration und der integrierten hierarchischen Informationsintegration vorgestellt. Da bei diesen Verfahren Attribute zu Konstrukten zusammengefasst werden, werden in Kapitel 5 Verfahren aufgezeigt, mit denen eine empirische Entwicklung solcher Konstrukte möglich ist. Ausgehend von den theoretischen Grundlagen erfolgt in Kapitel 6 die empirische Entwicklung von Konstrukten und in Kapitel 7 die Anwendung des Verfahrens der integrierten Informationsintegration im Regionalverkehr. Die Analyse dieser Experimente umfasst die Auswahl eines geeigneten Modells, die Überprüfung der Modellparameter in Bezug auf Unterschiede zwischen Nutzergruppen sowie die Überprüfung der hierarchischen Struktur des Modells. Im Ergebnis werden Nutzenfunktionen differenziert nach Nutzergruppen bestimmt und darauf aufbauend Zahlungsbereitschaften für Veränderungen von Ausprägungen der Qualitätsattribute ermittelt. Darüber hinaus wird anhand einer Simulation

exemplarisch veranschaulicht, wie die Ergebnisse eingesetzt werden können, um Auswirkungen auf die Verkehrsmittelwahl infolge einer Änderung der Qualität zu prognostizieren. Abschließend wird in Kapitel 8 ein Fazit gezogen und ein Ausblick auf weiteren Forschungsbedarf gegeben.

2 Qualität im Schienenpersonennahverkehr

2.1 Rahmenbedingungen im Schienenpersonennahverkehr

Den institutionellen Rahmen des Schienenpersonennahverkehrs bilden Gesetze auf europäischer, nationaler und regionaler Ebene. Diese Gesetze sind durchzogen von dem Kerngedanken, dass Verkehrsdienstleistungen im Bereich des öffentlichen Personennahverkehrs, die als Dienstleistungen von allgemeinem (wirtschaftlichem) Interesse bzw. als Aufgabe der Daseinsvorsorge angesehen werden, vielfach nicht wirtschaftlich betrieben werden können, sodass am Markt kein ausreichendes Angebot entstehen würde.

2.1.1 Rahmenbedingungen auf europäischer Ebene

Vor dem Hintergrund, dass am Markt kein ausreichendes Angebot entstehen würde, wurden in der am 3.12.2009 in Kraft getretenen Verordnung (EG) Nr. 1370/2007 die Rahmenbedingungen für einen regulierten Wettbewerb geschaffen und dabei versucht, die Unsicherheiten⁸ der durch sie aufgehobenen Verordnung (EWG) Nr. 1191/96 zu beseitigen. In der geltenden Verordnung wird festgelegt, wie die „zuständigen Behörden des öffentlichen Personenverkehrs“ im Einklang mit dem europäischen Gemeinschaftsrecht agieren können, „um die Erbringung von Dienstleistungen von allgemeinem Interesse zu gewährleisten, die unter anderem zahlreicher, sicherer, höherwertig oder preisgünstiger sind als diejenigen, die das freie Spiel des Marktes ermöglicht hätte.“⁹

Aufgrund des fehlenden wettbewerblichen Angebots werden gemeinwirtschaftliche Verkehrsdienstleistungen von der öffentlichen Hand vergeben. Dabei bedeutet gemeinwirtschaftlich, dass die Verkehrsdienstleistungen im allgemeinen Interesse liegen und von dem Betreiber ansonsten aufgrund seines eigenen wirtschaftlichen Interesses nicht, nicht im gleichen Umfang oder nicht zu den entsprechenden Bedingungen

⁸ Vgl. *Theobald/Kafka* (2003), S. 11–13.

⁹ Art. 1 Abs. 1 *Verordnung (EG) Nr. 1370/2007*.

angeboten würden.¹⁰ Als Vergabeform ist dabei neben der wettbewerblichen Vergabe auch, sofern das nationale Recht dem nicht im Wege steht, eine Direktvergabe oder Selbsterbringung durch die „zuständige örtliche Behörde“ vorgesehen.¹¹ Sofern dem ausgewählten Betreiber ausschließliche Rechte und/oder Ausgleichsleistungen zugestanden werden, ist ein öffentlicher Dienstleistungsauftrag erforderlich, welcher neben Art, Umfang und Dauer der Verkehrsdienstleistungen unter anderem auch die Aufteilung der entstehenden Kosten sowie die Aufteilung der Einnahmen aus dem Fahrkartenverkauf regelt. Diese Einnahmen können entweder vollständig von der zuständigen Behörde oder dem Betreiber vereinnahmt werden oder zwischen diesen beiden Parteien aufgeteilt werden.¹² Zudem ermöglicht die Verordnung, dass die zuständigen Behörden soziale Kriterien oder Qualitätskriterien vorgeben, die einer Aufrechterhaltung oder Erhöhung des Qualitätsstandards dienen.¹³ Näheres dazu ist gemäß dem Subsidiaritätsprinzip von den Mitgliedstaaten zu regeln.

2.1.2 Rahmenbedingungen auf nationaler Ebene

Den institutionellen Rahmen auf nationaler Ebene bildet im Wesentlichen das am 1.1.1994 in Kraft getretene Gesetz zur Neuordnung des Eisenbahnwesens (Eisenbahnneuordnungsgesetz - ENeuOG), mit dem die Bahnreform eingeleitet wurde. Bei diesem Gesetz handelt es sich um ein sogenanntes Artikelgesetz, d. h. es beinhaltet weitere Gesetze, darunter in Artikel 4 das Gesetz zur Regionalisierung des öffentlichen Personennahverkehrs (Regionalisierungsgesetz) und in Artikel 5 das Allgemeine Eisenbahngesetz (AEG).

Das Regionalisierungsgesetz, welches am 1.1.1996 in Kraft getreten ist, sieht die „Sicherstellung einer ausreichenden Bedienung der Bevölkerung mit Verkehrsleistungen im öffentlichen Personennahverkehr“ als eine Aufgabe der Daseinsvorsorge an.¹⁴ In diesem Gesetz, ebenso wie in dem am 1.1.1994 in Kraft getretenen Allgemeinen Eisenbahngesetz, wird der öffentliche Personennahverkehr definiert als die „allgemein zugängliche

¹⁰ Vgl. Art. 2 Abs. e) *Verordnung (EG) Nr. 1370/2007*.

¹¹ Vgl. Art. 5 *Verordnung (EG) Nr. 1370/2007*.

¹² Vgl. Art. 4 Abs. 2 *Verordnung (EG) Nr. 1370/2007*.

¹³ Vgl. Präambel Abs. 17 *Verordnung (EG) Nr. 1370/2007*.

¹⁴ Vgl. § 1 Abs. 1 *RegG*.

Beförderung von Personen mit Verkehrsmitteln im Linienverkehr, die überwiegend dazu bestimmt sind, die Verkehrsnachfrage im Stadt-, Vorort- oder Regionalverkehr zu befriedigen. Das ist im Zweifel der Fall, wenn in der Mehrzahl der Beförderungsfälle eines Verkehrsmittels die gesamte Reiseweite 50 Kilometer oder die gesamte Reisezeit eine Stunde nicht übersteigt.“¹⁵ Eine derartige Definition ist relativ unscharf. Darüber hinaus wird der Regionalverkehr nicht vom Stadt- und Vorortverkehr abgegrenzt.

In Deutschland lag bis zur Bahnreform die finanzielle und organisatorische Verantwortung des SPNV beim Bund.¹⁶ Durch das Regionalisierungsgesetz wurde die Erbringung gemeinwirtschaftlicher Verkehrsleistungen den Ländern übertragen¹⁷, wobei die für diese Aufgabe „zuständigen Stellen“ durch Landesrecht bestimmt werden.¹⁸ Im Gegenzug für die Aufgabenübernahme erhalten die Länder Gelder aus dem Mineralölsteueraufkommen des Bundes, mit denen „insbesondere“ der SPNV zu finanzieren ist.¹⁹ Diese als Regionalisierungsmittel bezeichneten Gelder sind dynamisiert, wobei die Höhe der Gelder ab dem Jahr 2015 neu festzulegen ist.²⁰

Im Allgemeinen Eisenbahngesetz befinden sich mitunter Regelungen bezüglich der anzustrebenden Wettbewerbsbedingungen, die sowohl den intra- als auch den intermodalen Wettbewerb betreffen: So soll das Gesetz „der Gewährleistung [...] eines attraktiven Verkehrsangebots auf der Schiene sowie der Sicherheit eines wirksamen und unverfälschten Wettbewerbs auf der Schiene beim Erbringen von Eisenbahnverkehrsleistungen [...]“ dienen.²¹ Zudem sollen die „Wettbewerbsbedingungen der Verkehrsträger angeglichen“ und eine „volkswirtschaftlich sinnvolle

¹⁵ § 2 *RegG*, § 2 Abs. 5 *AEGL*.

¹⁶ Vgl. *Laaser* (1994), S. 3.

¹⁷ Vgl. § 4 *RegG*.

¹⁸ Vgl. § 1 Abs. 2 und § 4 *RegG*.

¹⁹ Vgl. § 5 Abs. 1 und § 6 Abs. 1 *RegG*.

²⁰ Vgl. § 5 Abs. 5 *RegG*. Darüber hinaus erhalten die Länder Gelder nach § 1 *Entflechtungsgesetz* und nach § 1 *Gemeindeverkehrsfinanzierungsgesetz*. Es sei jedoch angemerkt, dass die jährliche Steigerung der Regionalisierungsmittel geringer ist als die Kostensteigerung (insbesondere Energiekosten, Personalkosten und Infrastruktur-entgelte), vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 318.

²¹ § 1 Abs. 1 *AEGL*.

Aufgabenteilung“ durch „einen lauterer Wettbewerb der Verkehrsträger“ ermöglicht werden.²²

2.1.3 Rahmenbedingungen auf regionaler Ebene

Den institutionellen Rahmen auf regionaler Ebene bilden die ÖPNV-Gesetze der Länder, welche im Zuge der Regionalisierung erlassen wurden. In diesen Gesetzen sind unter anderem die Zuständigkeiten für den SPNV geregelt.

Das Gesetz über den öffentlichen Personenverkehr in Nordrhein-Westfalen (ÖPNVG NRW) sieht den ÖPNV²³ ebenfalls als Aufgabe der Daseinsvorsorge an.²⁴ Aber auch lenkungspolitische Ziele werden verfolgt: Das Verkehrsangebot auf der Schiene soll verbessert und dadurch das Verkehrsaufkommen des SPNV gesteigert werden.²⁵ Es besteht der Grundsatz, landesweit eine „angemessene Bedienung der Bevölkerung durch den ÖPNV zu gewährleisten“. Dabei wird eine Verkehrsbedienung als „angemessen“ bezeichnet, „die den Bedürfnissen der Fahrgäste nach hoher Pünktlichkeit und Anschlusssicherheit, fahrgastfreundlich ausgestalteten, sicheren und sauberer Fahrzeugen sowie Stationen und Haltestellen, bequemen Zugang zu allen für den Fahrgast bedeutsamen Informationen, fahrgastfreundlichem Service und einer geeigneten Verknüpfung von Angeboten des ÖPNV mit dem motorisierten und nicht motorisierten Individualverkehr Rechnung trägt.“²⁶ Zudem soll die Attraktivität des ÖPNV mitunter durch einheitliche und nutzerfreundliche Tarife sowie durch einheitliche Qualitätsstandards gesteigert werden.²⁷

Bei der Sicherstellung eines „angemessenen“ ÖPNV wird der SPNV bevorzugt. Jedoch sollen in ländlichen Gegenden Busnetze ein Grundnetz

²² Vgl. § 1 Abs. 5 AEG.

²³ Die Definition des ÖPNV in diesem Gesetz entspricht den Definitionen des ÖPNV im *Regionalisierungsgesetz* und im *Allgemeinen Eisenbahngesetz*, vgl. § 1 Abs. 2 ÖPNVG NRW.

²⁴ Vgl. § 1 Abs. 1 ÖPNVG NRW.

²⁵ Vgl. § 2 Abs. 2 ÖPNVG NRW.

²⁶ § 2 Abs. 3 ÖPNVG NRW.

²⁷ Vgl. § 2 Abs. 3 ÖPNVG NRW.

des SPNV ergänzen; bei geringer Nachfrage sollen alternative Bedienungsformen wie beispielsweise Rufbusse eingesetzt werden.²⁸

Aufgabenträger für die Planung, Organisation und Ausgestaltung des SPNV sind die Kreise und kreisfreien Städte, welche die „zuständige“ Behörde im Sinne der Verordnung (EG) 1370/2007 darstellen.²⁹ Die Kreise, kreisfreien Städte und bestehenden Zweckverbände schließen sich zur gemeinsamen Aufgabenwahrnehmung zu einem Zweckverband oder einer gemeinsamen Anstalt zusammen.³⁰ Vom Land werden zur Aufgabenwahrnehmung Gelder zur Verfügung gestellt.³¹

Die Aufgaben des Zweckverbands bzw. der gemeinsamen Anstalt schließen auch die Aufstellung eines Nahverkehrsplans ein. Dieser hat zum Ziel, den ÖPNV zu sichern bzw. zu verbessern und die „öffentlichen Verkehrsinteressen des Nahverkehrs“ zu konkretisieren. Dazu zählen beispielsweise Vorgaben bezüglich der einzusetzenden Fahrzeuge sowie Mindestanforderungen von Betriebszeiten und Anschlussbeziehungen, die mit der „angemessenen“ Verkehrsbedienung verbunden sind.³²

2.1.4 Nahverkehrsplan des Zweckverbands Nahverkehr Westfalen-Lippe

Fünf regionale westfälische Zweckverbände haben sich zum 1.1.2008 zu einem neuen Zweckverband, dem Zweckverband Nahverkehr Westfalen-Lippe (NWL) zusammengeschlossen.³³ Der NWL ist einer von drei Aufgabenträgern für den SPNV in NRW nach dem ÖPNV-Gesetz NRW.³⁴ Als solcher ist er für die Aufstellung eines Nahverkehrsplans für den SPNV zuständig.

²⁸ Vgl. § 2 Abs. 5 und 7 ÖPNVG NRW.

²⁹ Vgl. § 3 ÖPNVG NRW.

³⁰ Vgl. §§ 5, 5a und 6 ÖPNVG NRW.

³¹ Vgl. § 11 ÖPNVG NRW, vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 309.

³² Vgl. § 8 Abs. 1 und 3 ÖPNVG NRW.

³³ Dabei handelt es sich um die Zweckverbände Nahverkehrsverbund Paderborn/Höxter (NPH), VerkehrsVerbund OstWestfalenLippe (VVOWL), Zweckverband Personennahverkehr Westfalen-Süd (ZWS), Zweckverband SPNV Münsterland (ZVM) und Zweckverband SPNV Ruhr-Lippe (ZRL), denen bis zum 31.12.2007 die Aufgabenträgerschaft für den SPNV oblag, vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 15.

Die Beschreibung des Nahverkehrsplans dieses Zweckbands ist durch die Auswahl der Erhebungsregion begründet.

³⁴ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 15 und § 5 ÖPNVG NRW.

Im Nahverkehrsplan des NWL ist ein Leitbild formuliert, welches darin besteht, den „Menschen in Westfalen-Lippe eine Mobilität zu ermöglichen, die ihren Bedürfnissen entspricht“. Dieses soll durch eine „Bereitstellung adäquater SPNV-Dienstleistungen unter Beachtung des Wirtschaftlichkeit und Sicherung der Qualität erreicht werden“.³⁵ Die zur Erfüllung dieses Leitbilds genannten Ziele bestehen unter anderem in der Berücksichtigung unterschiedlicher Nutzergruppen, einem integrierten Verkehrssystem, einer hohen Qualität und einer größtmöglichen Wirtschaftlichkeit durch effizienten Mitteleinsatz.³⁶

Dem Nahverkehrsplan liegt die Annahme zugrunde, dass ein Zusammenhang zwischen der Qualität des SPNV und der Kundenzufriedenheit besteht, und dass diese das Wiederwahl- und Weiterempfehlungsverhalten und die Abwanderungsresistenz beeinflusst.³⁷ Dabei wird auf Kundenzufriedenheitsmessungen verwiesen, in denen sich insbesondere „so genannte Schlüsselfaktoren wie Taktfrequenz, Verbindungen und Anschlüsse sowie die Schnelligkeit der Verkehrsverbindungen“ als besonders wichtig herausstellten. Mit dem Ziel einer (weiteren) Veränderung der Verkehrsmittelwahl soll daher einem integralen Fahrplan eine besondere Bedeutung beigemessen werden. Weiter wird darauf hingewiesen, dass auch Fahrzeug und Fahrzeugausstattung, Sitzplatzverfügbarkeit sowie Sauberkeit in Fahrzeugen und an Stationen zur Kundenzufriedenheit beitragen würden.³⁸

Im Nahverkehrsplan werden Grundstandards und anzustrebende Ziele hinsichtlich der Qualität und Quantität des Verkehrsangebots festgelegt. Die Festlegung erfolgte im Einklang mit bestehenden Vorschriften³⁹ und Empfehlungen und zumindest teilweise vor dem Hintergrund der bisherigen Fahrgastnachfrage bzw. der Bedeutung der Qualität für den bestehenden Kunden.

Die Vorgaben zur Bedienungshäufigkeit und zu Bedienzeiten hängen von der bisherigen Fahrgastnachfrage ab. Für einzelne Streckenabschnitte im Netz wurde ein durchschnittlicher Querschnitt (Reisenden-Kilometer je

³⁵ *Nahverkehr Westfalen-Lippe* (2011), S. 19.

³⁶ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 19–20.

³⁷ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 84.

³⁸ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 85.

³⁹ Hierzu zählen umwelt-, sicherheits- und sozialrelevante Gesetze und Regelwerke.

Kilometer Betriebslänge) berechnet, in Abhängigkeit dessen die Anzahl der angebotenen Verbindungen und die Bedienzeiten festgelegt werden.⁴⁰ Zudem bestehen Vorgaben bezüglich der Pünktlichkeit, einer Mindestservicequote (Anteil der Züge mit Begleitpersonal, wobei das Personal Aufgaben in Vertrieb, Sicherheit, Sauberkeit und Service wahrnehmen soll) und der Anschlusssicherheit.⁴¹

Die Anforderungen und Empfehlungen der einzusetzenden Fahrzeuge orientieren sich an den „Empfehlungen für Anforderungen an Fahrzeuge in Vergabeverfahren für Mitglieder der BAG-SPNV“⁴². Sie beziehen sich beispielsweise auf Kapazitätsanforderungen, Sitzplatzverfügbarkeit, Art und Anordnung der Sitzplätze, Gepäckablagen, Mindestgangbreiten, Durchgangsmöglichkeit in den Fahrzeugen, Klimatisierung der Fahrzeuge (Heizung, Lüftung und Kühlung), Toiletten, Informationsanzeigen im und am Fahrzeug (z. B. mit Angaben zu Richtungsverlauf, Störungen und aktuellen Anschlusszügen), Ein- und Ausstiegsmöglichkeiten und Videoüberwachung.⁴³

Aufgrund des Spannungsfelds beim Fahrkartenvertrieb zwischen möglichst geringen Kosten einerseits und der Kundenorientierung andererseits erfolgt die Aufrechterhaltung (bzw. Neu- oder Wiedereinrichtung) von personenbedientem Vertrieb in Abhängigkeit der bisherigen Ein- und Aussteigerzahlen an den jeweiligen Bahnhöfen.⁴⁴

Im Nahverkehrsplan des NWL wird in Bezug auf eine vollständige Harmonisierung des Tarifsystems im Sinne einer einheitlichen Preisgestaltung der Verkehrsverbünde darauf verwiesen, dass dazu „umfangreiche Vorarbeiten“ erforderlich seien, wobei als Beispiele unter anderem die Untersuchung der mengenmäßigen und finanziellen Auswirkungen infolge einer Tarifänderung genannt werden.⁴⁵ Dahingegen erscheinen die

⁴⁰ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 69–73.

⁴¹ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 84–85.

⁴² Vgl. *Arbeitskreis „Fahrzeuge“ der BAG-SPNV* (2006) und *Arbeitskreis „Fahrzeuge“ der BAG-SPNV* (2010). In diesen Empfehlungen der Bundesarbeitsgemeinschaft der Aufgabenträger des SPNV (BAG-SPNV) wird zwischen sogenannten Kann-, Soll- und Muss-Vorgaben unterschieden, die als Hilfestellung bei der Auswahl von Vorgaben dienen sollen.

⁴³ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 87–94.

⁴⁴ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 157–160.

⁴⁵ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 157.

Vorgaben bezüglich der Qualität des Verkehrsangebots eher ad hoc. Sofern Unterscheidungen bei den Qualitätsvorgaben erfolgen, orientieren sich diese an der bisherigen Nachfrage, welche nur aktuelle Fahrgäste, nicht aber potenzielle Fahrgäste einschließt. Mögliche Wechselwirkungen zwischen der bisherigen Qualität, beispielsweise hinsichtlich der Bedienungshäufigkeit, der Bedienzeiten oder dem personellem Fahrkartenverkauf und der Anzahl der Fahrgäste bleiben dabei unberücksichtigt.

Zudem werden im Nahverkehrsplan Anforderungen an das Netz und die Stationen dargelegt und die Erweiterung der Infrastruktur (neue Stationen, Reaktivierung von Strecken und Neubau von Streckenabschnitten) konkretisiert.⁴⁶

Der Nahverkehrsplan NWL dient als Grundlage für Wettbewerbsverfahren, bei denen die grundsätzlichen Anforderungen an das Verkehrsangebot, die Fahrzeuge und die Qualität bei der Ausschreibung weiter präzisiert werden.⁴⁷ Für den im Anschluss an das Wettbewerbsverfahren abzuschließenden Verkehrsvertrag verwendet der Zweckverband Nahverkehr Westfalen-Lippe die Vertragsform Bruttovertrag mit Anreizsystem.⁴⁸

Der Nahverkehrsplan sieht vor, dass im Verkehrsvertrag die Tatbestände Nichtleistung und Schlechtleistung zu berücksichtigen sind. Allgemein wird von einer Nichtleistung gesprochen, wenn eine (Haupt-)Leistung nicht erbracht wird, wohingegen eine Schlechtleistung bedeutet, dass die (Haupt-)Leistung nicht in der vertraglich vereinbarten Qualität erbracht wird.⁴⁹ Als Beispiele für Nichtleistungen werden im Nahverkehrsplan streckenbezogene Ausfälle, sitzplatzbezogene Ausfälle, fahrzeugbezogene Abweichungen und Abweichungen von der Servicequote genannt.⁵⁰

Zur Überprüfung der erbrachten Leistung in Bezug auf Pünktlichkeit, Vertriebs, Zustand und Sauberkeit der Fahrzeuge und Stationen, Fahrgastinformationen im Zug und an der Station, Zugbegleitpersonal und Kundendialog sollen geeignete Messverfahren eingesetzt oder sogenannte

⁴⁶ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 94–145.

⁴⁷ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 305.

⁴⁸ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 305–306. Vgl. zu den Vertragsformen Kap. 2.1.5.

⁴⁹ Vgl. *Bracher et al.* (2004), S. 137 und *Gorter et al.* (2000), S. 15.

⁵⁰ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 85.

Profitester-Erhebungen durchgeführt werden. Zudem soll die Messung der Kundenzufriedenheit in Bezug auf Fahrgastinformation, Zugbegleitung sowie Zustand und Sauberkeit der Fahrzeuge erfolgen.⁵¹

2.1.5 Vertragsformen und Anreizstrukturen

Allgemein können für die Gestaltung eines Verkehrsvertrags unterschiedliche Vertragsformen eingesetzt werden. Die Vertragsformen unterscheiden sich im Wesentlichen dahingehend, wie das Einnahmerisiko aus Fahrgeldeinnahmen zwischen dem Aufgabenträger und dem Verkehrsunternehmen verteilt ist und damit einhergehend, welche Anreizstrukturen bestehen. Als Vertragsformen können der Bruttovertrag einerseits und der Nettovertrag andererseits unterschieden werden, sowie eine Kombination daraus, die als Mischvertrag, Brutto-Anreizvertrag oder Anreizvertrag bezeichnet wird.

Beim Bruttovertrag liegt das Einnahmerisiko aus Fahrgeldeinnahmen beim Aufgabenträger. Das Verkehrsunternehmen erhält für die Erbringung der Verkehrsleistung einen festen Geldbetrag vom Aufgabenträger. In diesem Betrag sind die Fahrgeldeinnahmen bereits enthalten. Für das Verkehrsunternehmen bietet diese Vertragsform den Vorteil der hohen Planungssicherheit. Allerdings besteht kein Anreiz, viele Fahrgäste zu befördern sowie Maßnahmen zur Qualitätserhaltung oder Qualitätssteigerung durchzuführen.⁵² Es ist somit erforderlich, dass der Aufgabenträger das gewünschte Qualitätsniveau schon bei der Ausschreibung sehr detailliert definiert, anschließend vertraglich festlegt und dessen Erfüllung kontrolliert. Bei dieser Vertragsform wird das Verkehrsunternehmen das Handeln ausschließlich an den Forderungen des Aufgabenträgers ausrichten, sodass der Aufgabenträger und nicht der Fahrgast der eigentliche Kunde ist.⁵³

Beim Nettovertrag liegt das Einnahmerisiko aus Fahrgeldeinnahmen beim Verkehrsunternehmen. Da die Fahrgeldeinnahmen nicht in dem (geringeren) festen Geldbetrag enthalten sind, den das Verkehrsunternehmen vom Aufgabenträger erhält, sondern darüber hinaus einbehalten bzw. gezahlt

⁵¹ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 86.

⁵² Vgl. *Bracher et al.* (2004), S. 114.

⁵³ Vgl. *Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006), S. 6.

werden, besteht für das Verkehrsunternehmen ein Anreiz, viele Fahrgäste zu befördern.⁵⁴ Eine detaillierte Vorgabe eines gewünschten Qualitätsniveaus durch den Aufgabenträger entfällt bei dieser Vertragsform.⁵⁵ Das Verkehrsunternehmen wird ökonomisch sinnvolle Maßnahmen zur Qualitätserhaltung oder Qualitätssteigerung durchführen. Dabei ist eine Maßnahme für ein Verkehrsunternehmen ökonomisch sinnvoll, wenn die zusätzlichen Fahrgeldeinnahmen größer sind als die Kosten zur Durchführung der Qualitätsmaßnahme. Ökonomisch nicht sinnvolle Maßnahmen wird das Verkehrsunternehmen hingegen nicht durchführen, was ein geringeres Qualitätsniveau zur Folge hätte.⁵⁶ Aufgrund dieser Anreizstruktur richtet das Verkehrsunternehmen sein Handeln am Fahrgast und nicht am Aufgabenträger aus. Für das Verkehrsunternehmen besteht das Problem einer geringen Planungssicherheit, denn es kann nicht alle Einflüsse auf das Wahlverhalten der Fahrgäste kontrollieren.⁵⁷ Beispielsweise könnte ein schlechtes Busangebot dazu führen, dass Fahrgäste die Bahn nicht oder nur schlecht erreichen können und daher den Pkw verwenden, obwohl das Bahnangebot optimal ist.

Beim Mischvertrag werden dem Bruttovertrag Bonus- bzw. Malus-Regelungen hinzugefügt, d. h. es erfolgen Abzüge bei Nicht- oder Schlechtleistung bzw. Zusatzzahlungen bei einer Übererfüllung der vereinbarten Leistung.⁵⁸ Die Bonus- bzw. Malus-Regelungen können sich dabei entweder auf ein vorab definiertes Qualitätsniveau oder auf eine zu erreichende Fahrgastzahl beziehen. Dadurch ergibt sich für das Verkehrsunternehmen ein Anreiz, das jeweilige Qualitätsniveau bzw. die Fahrgastzahl zu erreichen oder zu überschreiten.⁵⁹ Gleichzeitig besteht im Vergleich zum Nettovertrag für das Verkehrsunternehmen eine höhere Planungssicherheit. Zudem ermöglicht der Mischvertrag auch die Erreichung eines Qualitätsniveaus, welches bei einem reinen Nettovertrag für das Verkehrsunternehmen ökonomisch nicht sinnvoll wäre, aber beispielsweise zur Sicherstellung eines sozialen Mindeststandards oder aus

⁵⁴ Vgl. *Bracher et al.* (2004), S. 114 und *Eichmann et al.* (2006), S. 215.

⁵⁵ Vgl. *Gorter et al.* (2000), S. 17.

⁵⁶ Vgl. *Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006), S. 18.

⁵⁷ Vgl. *Eichmann et al.* (2006), S. 216.

⁵⁸ Vgl. *Gorter et al.* (2000), S. 7–8.

⁵⁹ Vgl. *Bracher et al.* (2004), S. 114–115.

lenkungspolitischen Gründen gewünscht ist.⁶⁰ Bei der Festlegung der Höhe der Bonus- und Malus-Zahlungen ist jedoch zu beachten, dass diese auch tatsächlich einen Anreiz für das Verkehrsunternehmen darstellen. Dies bedeutet, dass die Kosten zur Erbringung des jeweiligen Qualitätsniveaus ggf. abzüglich möglicher Bonus-Zahlungen nicht größer sein dürfen als die Kosten der Nichterbringung ggf. zuzüglich möglicher Malus-Zahlungen.⁶¹

Somit ist je nach verwendeter Vertragsform die Kenntnis des Zusammenhangs zwischen der Qualität und der Verkehrsmittelwahl für den Aufgabenträger und/oder für das Verkehrsunternehmen von Bedeutung. Bei einem Nettovertrag kann das Verkehrsunternehmen bei der Kenntnis dieses Zusammenhangs besser abschätzen, welche Maßnahmen zur Qualitätserhaltung oder Qualitätssteigerung ökonomisch sinnvoll sind und daran das Verkehrsangebot ausrichten, d. h. die Kenntnis des Zusammenhangs ist aus einzelwirtschaftlicher Sicht von Bedeutung. Den Aufgabenträger könnte die Kenntnis dieses Zusammenhangs und des zu erwartenden Qualitätsniveaus gegebenenfalls dazu veranlassen, stattdessen eine andere Vertragsform zu wählen, bei der ein Mindeststandard festgelegt wird.

Beim Bruttovertrag muss die gewünschte Qualität detailliert vorgegeben und vertraglich vereinbart werden, da sie sonst nicht vom Verkehrsunternehmen erbracht würde. Vor dem Hintergrund des effizienten Einsatzes öffentlicher Gelder sollte jedoch bei der Festlegung des Qualitätsniveaus darauf geachtet werden, dass gegebenenfalls angestrebte lenkungspolitische oder andere Zielsetzungen auch tatsächlich durch ein höheres Qualitätsniveau erreicht werden.⁶² Somit besteht ein gemeinwirtschaftliches Interesse an der Kenntnis des Zusammenhangs zwischen Qualität und Verkehrsmittelwahl. Dieses Interesse besteht auch beim Mischvertrag, um aufbauend auf einem quantifizierten Zusammenhang die Höhe der anreizkompatiblen Bonus- und Malus-Zahlungen festzulegen.

⁶⁰ Vgl. *Lehmbrock et al.* (2006), S. 59.

⁶¹ Vgl. *Bracher et al.* (2004), S. 135 und *Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006), S. 16.

⁶² Vgl. *Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006), S. 15.

2.2 Qualität und Kundenzufriedenheit

Sofern ein zu erbringendes Qualitätsniveau in einem Verkehrsvertrag vereinbart ist und für das Verkehrsunternehmen ansonsten kein Anreiz zu dessen Erbringung besteht, sollte ein geeignetes Verfahren zur Überprüfung, ob das Qualitätsniveau mindestens erbracht wird, eingesetzt werden. Eine Orientierungshilfe dafür soll die DIN EN 13816 darstellen, auf die im Folgenden eingegangen wird.

2.2.1 DIN EN 13816

Die Europäische Norm 13816, die gleichzeitig den Status einer Deutschen Norm hat, dient hauptsächlich dem Zweck, die „Qualitätsphilosophie für öffentliche Verkehre zu fördern sowie das Augenmerk auf die Bedürfnisse und Erwartungen der Kunden zu lenken [...]“.⁶³

In der Norm findet sich auch der Gedanke des effizienten Einsatzes öffentlicher Mittel wieder. So soll die Erfüllung der Norm dazu beitragen, die „zur Verfügung stehenden Ressourcen jenen Aufgaben zuzuordnen, welche mit größter Wahrscheinlichkeit zu zusätzlicher Kundenzufriedenheit und weiteren Einkünften bei den ÖPV-Beteiligten führen“.⁶⁴ Zudem werden Empfehlungen für die Messung der Dienstleistungsqualität gegeben, wobei Dienstleistungsqualität definiert ist als eine „Reihe von Qualitätskriterien und geeigneten Maßnahmen, für die der Dienstleistungsanbieter (derjenige, der Anspruch erhebt, die Norm zu erfüllen) verantwortlich ist“.⁶⁵

Als Konzept liegt der Norm der Qualitätskreis⁶⁶ zugrunde, welcher in Abb. 2.1 veranschaulicht ist. Der Qualitätskreis unterteilt sich in die Sicht

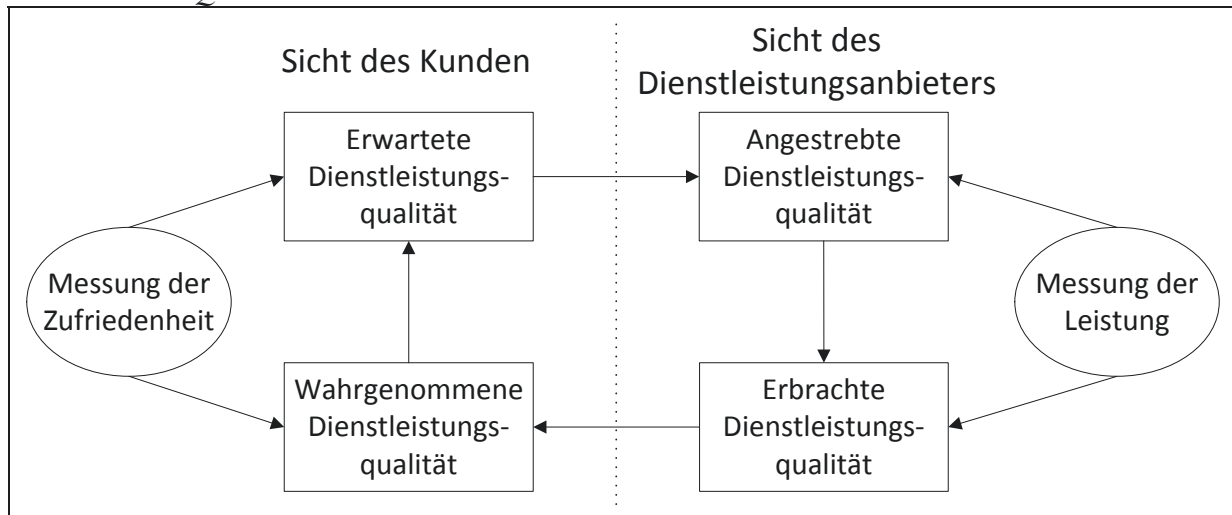
⁶³ DIN EN 13816, S. 4.

⁶⁴ DIN EN 13816, S. 4.

⁶⁵ Vgl. DIN EN 13816, S. 6. Aufgrund dieser Definition werden im Folgenden im Zusammenhang mit dieser Norm Qualitätsmerkmale als „Qualitätskriterien“ bezeichnet.

⁶⁶ Die beim Qualitätskreis vorgenommene Messung der Zufriedenheit basiert auf dem Erklärungsansatz des C/D-Paradigmas (Confirmation/Disconfirmation), bei dem ein Vergleich zwischen der wahrgenommenen Leistung (Ist-Leistung) und der erwarteten Leistung (Soll-Leistung) erfolgt. Entspricht die wahrgenommene Leistung mindestens der erwarteten Leistung, stellt sich Zufriedenheit ein, anderenfalls Unzufriedenheit, vgl. Homburg/Stock-Homburg (2006), S. 20–23, Giering (2000), S. 8–10 und Koschate (2002), S. 12–13. Dieser Vergleichsprozess ist auch Teil des sogenannten Gap-Modells

Abb. 2.1: Qualitätskreis



Quelle: DIN EN 13816, S. 6.

des Dienstleistungsanbieters und in die Sicht des Kunden. Aus Sicht des Dienstleistungsanbieters werden die angestrebte und die erbrachte Dienstleistungsqualität unterschieden, deren Vergleich bei der Leistungsmessung eine Rolle spielt. Aus Sicht des Kunden werden die wahrgenommene und die erbrachte Dienstleistungsqualität unterschieden. Die Abweichung zwischen diesen beiden Qualitäten ist Gegenstand der Kundenzufriedenheitsmessung.

In der Norm werden acht Kategorien angeführt, die sich auf die Wahrnehmung aus Kundensicht beziehen und gemeinsam die „Gesamtqualität“ bilden. Dabei handelt es sich um die Kategorien Verfügbarkeit, Zugänglichkeit, Informationen, Zeit, Kundenbetreuung, Komfort, Sicherheit und Umwelteinflüsse.⁶⁷ Diese Kriterien werden weiter untergliedert (Anhang A der Norm). Beispielsweise wird der Komfort in die Unterkategorien Benutzbarkeit von Fahrgasteinrichtungen, Raumangebot, Fahrkomfort, das Umfeld betreffende Bedingungen, zusätzliche Einrichtungen sowie Ergonomie aufgeteilt; bei der Benutzbarkeit von Fahrgasteinrichtungen wird wiederum zwischen Haltestellen und Fahrzeugen unterschieden. Die Norm sieht vor, dass grundsätzlich alle acht Kategorien zur Beschreibung der Qualität verwendet werden sollen, wenn möglich sogar mit den jeweiligen Unterkategorien.

der Kundenzufriedenheit, vgl. Parasuraman/Zeithaml/Berry (1985), S. 44–46 und Transportation Research Board (1999), S. 6–8.

⁶⁷ Vgl. DIN EN 13816, S. 8.

Zur Überprüfung der Qualität der erbrachten Dienstleistungen in Bezug auf die einzelnen Kriterien sollen entweder „geeignete Methoden“ (Anhang C der Norm) oder Alternativmethoden, die zu gleichwertigen Ergebnissen führen, verwendet werden.⁶⁸ Dabei werden als Verfahren zur Messung der Leistung bzw. der Zufriedenheit die Beurteilung durch Testkunden (*Mystery Shopping Surveys*) sowie die unmittelbare Leistungsmessung (*Direct Performance Measures*) bzw. Umfragen zur Kundenzufriedenheit (*Customer Satisfaction Surveys*) genannt. Bei den Messungen sollen nicht nur aktuelle Kunden, sondern auch potentielle Kunden berücksichtigt werden, welches dadurch begründet wird, dass letztere möglicherweise die Verkehrsdienstleistung aufgrund der derzeitigen Qualität nicht nutzen.⁶⁹

Für diese drei Verfahren erfolgt die Messung hinsichtlich einzelner Kriterien bzw. Unterkriterien. Obgleich bei der Messung der Kundenzufriedenheit darauf hingewiesen wird, dass die wichtigsten Kriterien vorrangig beurteilt werden sollen, erfolgt hierbei kein Hinweis zur Ermittlung der Wichtigkeit. Allerdings wird im Zusammenhang mit der Qualitätsstufe der erwarteten Dienstleistungsqualität, die als Summe der gewichteten Qualitätskriterien definiert ist, darauf hingewiesen, dass die Gewichtung auf einer qualitativen⁷⁰ Analyse basieren kann.⁷¹

2.2.2 Verfahren zur Messung der Qualität und der Kundenzufriedenheit

Zur Messung der Qualität oder der Zufriedenheit existiert eine Vielzahl von Verfahren. Diese können zunächst in anbieterbezogene (d. h. unternehmens- und mitarbeiterbezogene) und nachfragerbezogene Verfahren unterschieden werden. Eine Systematisierung der nachfragerbezogenen Verfahren ist in Abb. 2.2 dargestellt. Dabei wird in Bezug auf

- den Grad an Differenzierung,
- den Grad an Objektivität,

⁶⁸ Vgl. *DIN EN 13816*, S. 10.

⁶⁹ Vgl. *DIN EN 13816*, S. 22–23.

⁷⁰ Warum dabei nicht auf eine quantitative Analyse verwiesen wird, geht aus der Norm nicht hervor.

⁷¹ Vgl. *DIN EN 13816*, S. 7.

- die Ausdrücklichkeit und
- den zugrunde liegenden Messansatz

unterschieden.⁷² Die bei dieser Systematisierung nach *Schnippe* (2000, S. 162) aufgeführten Verfahrensbeispiele sind allerdings nicht frei von Überschneidungen: Beispielsweise kann im Rahmen eines Nationalen Barometers eine kompensatorische Messung eingesetzt werden oder eine Messung nach dem Willingness-to-Pay-Ansatz kann anhand eines dekompositionellen Verfahrens erfolgen.

Ogleich viele unterschiedliche Verfahren zur Messung der Qualität bzw. der Kundenzufriedenheit im ÖPNV eingesetzt werden, werden schwerpunktmäßig merkmalsorientierte Verfahren verwendet.⁷³

Bei der Messung der Zufriedenheit durch merkmalsorientierte Verfahren kann unterschieden werden, ob die Zufriedenheit direkt auf einer entsprechenden Skala⁷⁴ abgefragt wird oder ob erwartete und wahrgenommene Leistung separat erfragt werden und daraus die Zufriedenheit abgeleitet wird. Dabei bezieht sich die Zufriedenheit entweder auf eine bestimmte Dienstleistung bzw. ein bestimmtes Produkt (transaktionsspezifische Kundenzufriedenheit) oder auch auf bisherige Erfahrungen mit der Dienstleistung/dem Produkt (kumulative Kundenzufriedenheit).⁷⁵ Zudem können merkmalsorientierte Verfahren dahingehend unterschieden werden, ob die Wichtigkeit der einzelnen Qualitätsmerkmale direkt abgefragt oder indirekt ermittelt wird.

Eine gängige Vorgehensweise zur Messung der Kundenzufriedenheit besteht darin, dass die Qualitätsmerkmale jeweils einzeln auf einer Skala

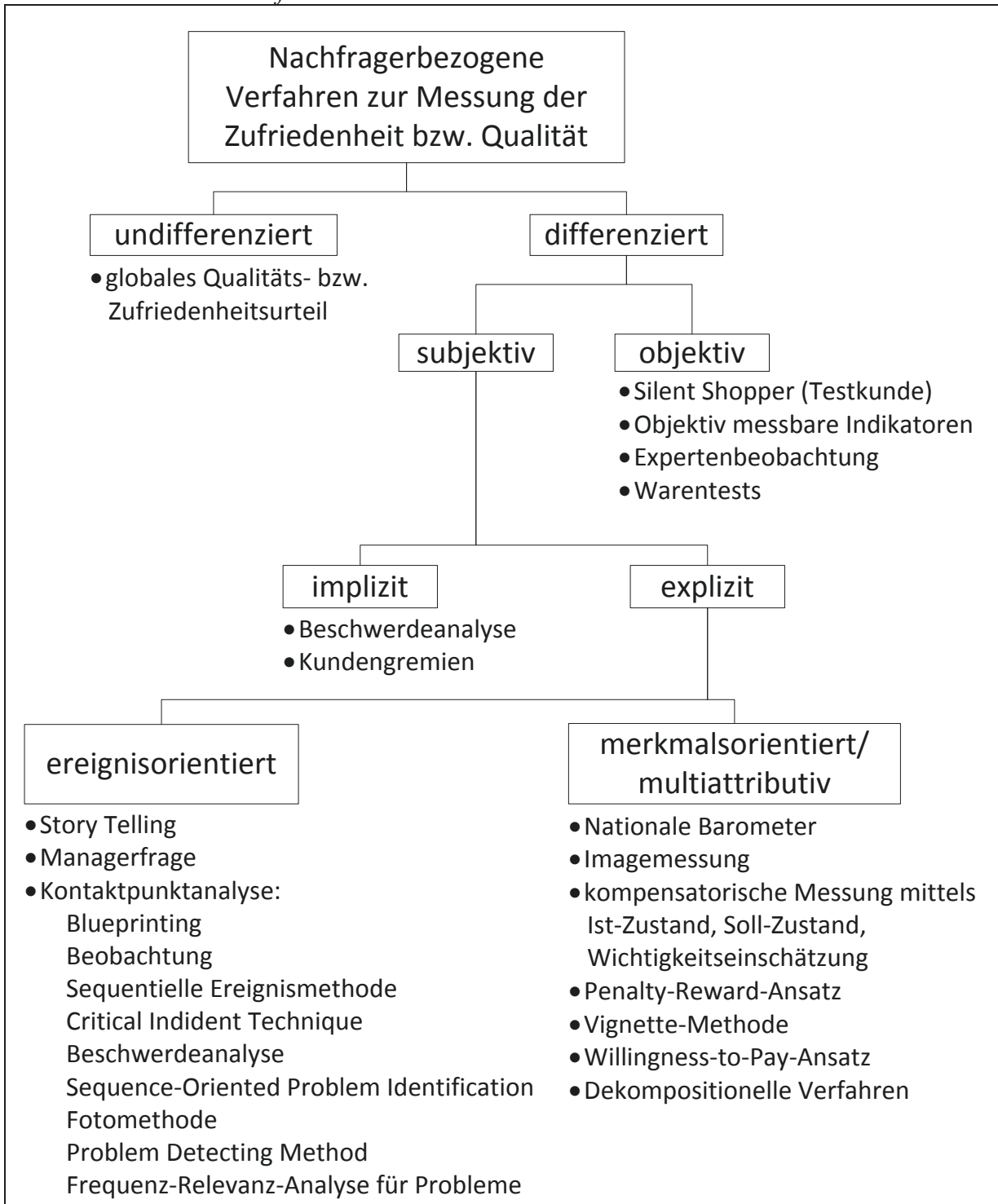
⁷² Vgl. *Schnippe* (2000), S. 163.

⁷³ Für eine Beschreibung der unterschiedlichen Verfahren vgl. beispielsweise *Schnippe* (2000), S. 174–202 und *Meffert/Bruhn* (2003), S. 288–326. Für Beispiele von Anwendungen der objektiven Verfahren vgl. *Felscher-Suhr/Meyer/Legath* (2011), *Plath* (2009) und *Lawerentz/Balling/Fleischer* (2002); für Beispiele von Anwendungen der impliziten Verfahren vgl. *Stiefelbusch/Fliegel* (2006) und *Panitz/Jobke/Jensch* (2012); für Beispiele von Anwendungen der ereignisorientierten Verfahren vgl. *Probst/Kahrs/Rieschick-Dziabas* (2001), *Karopka et al.* (2000) und *Plantsch* (2011); für Beispiele von Anwendungen der merkmalsorientierten Verfahren vgl. *TNS Infratest* (2012a), *Krietemeyer/Wergles* (2006), *Angermüller* (2007) und *Georgi/Wüst-Rocktäschel/Heller* (2006).

⁷⁴ Für eine Übersicht möglicher Skalen vgl. *Matzler/Bailom* (2002), S. 231–234.

⁷⁵ Vgl. *Giering* (2000), S. 11–14.

Abb. 2.2: Systematisierung nachfragerbezogener Verfahren zur Messung der Qualität bzw. der Zufriedenheit



Quelle: Schnippe (2000, S. 162).

in Bezug auf die Zufriedenheit sowie in Bezug auf die Wichtigkeit zu bewerten sind. Bei der direkten Abfrage der Wichtigkeit auf einer Skala besteht allerdings die Tendenz, dass viele Qualitätsmerkmale als sehr wichtig beurteilt werden, sodass keine deutliche Differenzierung möglich ist. Diese Problematik könnte durch die Bildung einer Rangfolge oder durch die Verwendung eines Konstantsummenverfahrens umgangen werden, bei dem eine feste Anzahl von Punkten auf die einzelnen Qualitätsmerkmale verteilt werden, um die relative Wichtigkeit auszudrücken.⁷⁶ Derartige Messverfahren erlauben keine Berücksichtigung von Interaktionen zwischen unterschiedlichen Merkmalen.

Eine weitere gängige Vorgehensweise besteht darin, neben der Zufriedenheit hinsichtlich einzelner Qualitätsmerkmale auch die Zufriedenheit insgesamt abzufragen, welche als Gesamt- oder Globalzufriedenheit bezeichnet wird. Unterschiedliche Verfahren können eingesetzt werden, um aus den Werten der Globalzufriedenheit und der Zufriedenheit hinsichtlich der einzelnen Qualitätsmerkmale deren jeweilige Wichtigkeit zu ermitteln.

Dazu zählt die Bestimmung von Korrelationen zwischen den einzelnen Qualitätsmerkmalen und der Globalzufriedenheit. Einerseits ist dieses Verfahren leicht anzuwenden, andererseits weisen oftmals viele Qualitätsmerkmale gleiche Korrelationen auf, sodass keine Differenzierung möglich ist.

Eine andere Auswertungsmöglichkeit stellt die multiple lineare Regressionsanalyse dar, die den Vorteil bietet, dass Interaktionen berücksichtigt werden können. Diese Vorgehensweise wird beispielsweise beim sogenannten ÖPNV-Kundenbarometer⁷⁷ verwendet. Sofern jedoch die Zufriedenheiten in Bezug auf die einzelnen Qualitätsmerkmale miteinander korrelieren (Multikollinearität), lässt sich der Einfluss der einzelnen

⁷⁶ Vgl. Matzler/Bailom (2002), S. 235–236, Helm/Steiner (2008), S. 248–249, Berekoven/Eckert/Ellenrieder (2004), S. 78, Kroeber-Riel/Weinberg/Gröppel-Klein (2009), S. 221 und Matzler/Sauerwein/Stark (2002), S. 282.

⁷⁷ Das ÖPNV-Kundenbarometer ist ein Messinstrument von TNS Infratest. Es dient zur Messung der Kundenzufriedenheit im ÖPNV, wobei einerseits die Globalzufriedenheit und andererseits die Zufriedenheit in Bezug auf einzelne Qualitätsmerkmale gemessen werden. Der Vergleich mit der Kundenzufriedenheit von anderen Verkehrsunternehmen bzw. Verkehrsverbänden soll gegebenenfalls als Ansatzpunkt für eine Verbesserung des Angebots dienen, vgl. TNS Infratest (2012a), S. 1–2 und TNS Infratest (2012b), S. 2–3.

Zufriedenheiten auf die Globalzufriedenheit nicht mehr eindeutig zuordnen. Diese Problematik kann umgangen werden, indem zunächst in einer Faktorenanalyse Qualitätsmerkmale, die miteinander korrelierende Zufriedenheiten aufweisen, zu sogenannten Faktoren zusammengefasst werden, die voneinander unabhängig sind. Diese Faktoren können anschließend in einer multiplen linearen Regressionsanalyse verwendet und deren Einfluss auf die Globalzufriedenheit bestimmt werden. Allerdings ergibt sich durch das Zusammenfassen zu Faktoren ein Informationsverlust.⁷⁸

Aufbauend auf den Zufriedenheiten und den relativen Wichtigkeiten kann ein Streudiagramm erstellt werden, welches als Zufriedenheitsportfolio, Zufriedenheitsprofil oder Handlungsportfolio bezeichnet wird.⁷⁹ In diesem Diagramm wird die Position der einzelnen Qualitätsmerkmale durch den jeweiligen Wert der Zufriedenheit auf der einen Achse und durch die relative Wichtigkeit auf der anderen Achse bestimmt. Es erfolgt oftmals eine Unterteilung der Achsen dergestalt, dass sich vier Felder ergeben. Für diese Unterteilung sind Schwellenwerte vorzugeben, die beispielsweise dem Median der Zufriedenheit bzw. dem Median der Wichtigkeit entsprechen. Zufriedenheiten oberhalb des Medians stehen für eine hohe Zufriedenheit und Werte unterhalb des Medians für eine geringe Zufriedenheit. Ebenso steht eine relative Wichtigkeit oberhalb des Medians für eine hohe Wichtigkeit und eine relative Wichtigkeit unterhalb des Medians für eine geringere Wichtigkeit. Darauf aufbauend soll der Handlungsbedarf abgeleitet werden:

- Qualitätsmerkmale mit hoher Wichtigkeit und hoher Zufriedenheit stellen bedeutsame Stärken dar, die keinen Handlungsbedarf erfordern.
- Qualitätsmerkmale mit hoher Wichtigkeit und geringer Zufriedenheit stellen bedeutsame Schwächen dar, die einen besonderen Handlungsbedarf erfordern.
- Qualitätsmerkmale mit geringer Wichtigkeit und geringer Zufriedenheit stellen unbedeutende Schwächen dar, die in Kauf genommen werden können.

⁷⁸ Vgl. *Transportation Research Board* (1999), S. 15–16.

⁷⁹ Eine derartige Anwendung erfolgt beispielsweise beim ÖPNV-Kundenbarometer, vgl. *TNS Infratest* (2012b), S. 3 und *Isfort* (2000), S. 41.

- Qualitätsmerkmale mit hoher Zufriedenheit und geringer Wichtigkeit stellen unbedeutende Stärken dar, die einen Zusatznutzen bieten, als Selbstverständlichkeit wahrgenommen werden oder irrelevant sind.⁸⁰

Allerdings sind mit einer derartigen Bestimmung des Handlungsbedarfs einige Nachteile verbunden. Zum einen erfolgt die Festlegung der Schwellen eher willkürlich. Über die Zuordnung zu den einzelnen Feldern hinaus erfolgt keine weitere Differenzierung, d. h. für alle Qualitätsmerkmale innerhalb eines Felds wird der gleiche Handlungsbedarf angesehen und die jeweiligen Werte der Zufriedenheit bzw. der relativen Bedeutung werden nicht weiter berücksichtigt.⁸¹ Eine etwas stärkere Differenzierung wäre möglich, wenn die Anzahl der Felder erhöht wird, sodass beispielsweise neun statt vier Felder verwendet werden. Doch auch bei dieser weiteren Differenzierung besteht das Problem, dass unterstellt werden muss, dass ein Zusammenhang zwischen der Kundenzufriedenheit und der Verkehrsmittelwahl bzw. der Verkehrsmittelnutzung besteht. Zudem werden bei einer derartigen Bestimmung des Handlungsbedarfs die Kosten der Maßnahmen zur Erhöhung der Kundenzufriedenheit nicht berücksichtigt.

Allgemein liegt bei derartigen Verfahren der Schwerpunkt auf der Messung der Leistung bzw. der Zufriedenheit; ein Zusammenhang zur Verkehrsmittelwahl besteht nicht bzw. nicht direkt. Somit bleibt offen, inwiefern die Erfüllung der oben beschriebenen Norm in Form einer hohen Qualität bzw. einer hohen Kundenzufriedenheit tatsächlich dazu beiträgt, die „zur Verfügung stehenden Ressourcen jenen Aufgaben zuzuordnen, welche mit größter Wahrscheinlichkeit zu zusätzlicher Kundenzufriedenheit und weiteren Einkünften bei den ÖPV-Beteiligten führen“.⁸²

Im Folgenden werden exemplarisch drei sehr unterschiedliche Studien mit einem Bezug zur Verkehrsmittelwahl bzw. zum wirtschaftlichen Erfolg vorgestellt und kritisch gewürdigt.

⁸⁰ Vgl. Schnippe (2000), S. 198, Beutin (2006), S. 158–159, *Transportation Research Board* (1999), S. 17 und Engehausen (1999), S. 210–211.

⁸¹ Vgl. *Transportation Research Board* (1999), S. 17.

⁸² *DIN EN 13816*, S. 4.

2.2.3 Kundenzufriedenheit und wirtschaftlicher Erfolg

Bäumer/Hussinger/Pfeiffer (2010) untersuchten, inwiefern ein Zusammenhang zwischen der Kundenzufriedenheit und dem wirtschaftlichem Erfolg des Verkehrsunternehmens bzw. des Verkehrsverbundes besteht. Die Studie gliederte sich im Wesentlichen in zwei Analysebereiche. Zum einen wurde untersucht, inwiefern qualitätsverändernde Maßnahmen die Zufriedenheit in Bezug auf einzelne Qualitätsmerkmale und die Globalzufriedenheit beeinflussen, und zum anderen, ob ein Zusammenhang zwischen Globalzufriedenheit und Kennziffern des wirtschaftlichen Erfolgs besteht.

Zur Analyse des ersten Zusammenhangs war eine Vorher-/Nachher-Erhebung konzipiert, sodass Verkehrsunternehmen bzw. Verkehrsverbände, in denen angebotsverändernde Maßnahmen umgesetzt werden sollten, identifiziert und davon drei Verkehrsunternehmen als Fallbeispiele (SWB Stadtwerke Bonn Verkehrs GmbH, RNV Rhein-Neckar-Verkehrs GmbH und SVT Stadtverkehr Tübingen) ausgewählt wurden. Bei den durchgeführten Maßnahmen handelte es sich um eine Neuordnung des Busnetzes, die im Wesentlichen Änderungen des Liniennetzes und der Bedienungshäufigkeit zur Folge hatte, um eine Neugestaltung des ÖPNV-Linienkonzepts, die zu einer Änderung der Buslinien und zum Wegfall einer Tramlinie führte, bzw. um eine Einführung eines dynamischen Fahrgast-Informationssystems⁸³ und um die Anschaffung neuer Fahrzeuge, die das Platzangebot vergrößern und durch beschleunigtes Ein- und Aussteigen (aufgrund einer zusätzlichen Tür) die Pünktlichkeit verbessert sollte.⁸⁴

Die Datenerhebung erfolgte als computergestützte telefonische Befragung. Je nach Nutzungshäufigkeit des ÖPNV wurden die Befragten drei Teilstichproben zugeordnet, die hier vereinfachend als regelmäßige Nutzer, Selten-Nutzer und Nicht-Nutzer bezeichnet werden. Der Befragungs-

⁸³ Anders als geplant befand sich bei der Nachher-Erhebung das dynamische Fahrplan-Informationssystem noch im Testbetrieb. Zudem erfolgte auch eine Veränderung der Rahmenbedingungen im Untersuchungsgebiet beispielsweise in Form einer kostenlosen ÖPNV-Nutzung an Samstagen und der Beeinträchtigung des Verkehrs durch eine Großbaustelle, die Verlängerungen der Reisezeit zur Folge hatten.

⁸⁴ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 51–57 und *Bäumer et al.* (2009), S. 24–26.

ablauf war abhängig von der Zugehörigkeit zu diesen drei Teilstichproben.

Bei den regelmäßigen Nutzern bezog sich die Befragung auf eine konkrete Fahrt des Befragten: Für diese Fahrt wurden die Globalzufriedenheit und die Zufriedenheit in Bezug auf elf Qualitätsmerkmale⁸⁵ (Einzelzufriedenheiten) erhoben. Unabhängig davon wurde die allgemeine Globalzufriedenheit und die Wichtigkeit der elf Qualitätsmerkmale erfragt. Als Skala der (Global-)Zufriedenheit sowie der Wichtigkeit wurde eine fünfstufige Skala verwendet (1 = sehr zufrieden, sehr gut bzw. sehr wichtig, 5 = gar nicht zufrieden, gar nicht gut bzw. gar nicht wichtig). Bei den Selten-Nutzern erfolgte eine allgemeine Abfrage der Globalzufriedenheit und der Zufriedenheit in Bezug auf die elf Qualitätsmerkmale und deren Wichtigkeit. Zudem wurden die regelmäßigen Nutzer und Selten-Nutzer bezüglich des Weiterempfehlungsverhaltens, der Wieder- und Zusatzkaufabsicht und der Weiterempfehlungsabsicht befragt. Bei den Nicht-Nutzern erfolgte eine Erhebung der grundsätzlichen Einstellung gegenüber dem ÖPNV und der Gründe für die Nicht-Nutzung. Darüber hinaus wurden unabhängig vom Nutzungsverhalten soziodemografische Daten und Angaben zu sogenannten Critical Incidents erhoben, d. h. besonders gute und/oder besonders schlechte Erlebnisse im ÖPNV. Nach einem Jahr wurden dieselben Personen auf die gleiche Weise erneut befragt, wobei aufgrund eines geänderten Nutzungsverhaltens die Zuordnung zu den Teilstichproben variieren konnte.⁸⁶

Mehrere Analysen wurden durchgeführt, um den Zusammenhang zwischen der Zufriedenheit in Bezug auf die einzelnen Qualitätsmerkmale und die Globalzufriedenheit zu analysieren: Ein Zufriedenheitsportfolio wurde anhand der Mittelwerte der elf Einzelzufriedenheiten und der direkt erhobenen Wichtigkeiten erstellt. Darüber hinaus wurden die Wichtigkeiten der einzelnen Merkmale im Rahmen einer multiplen linearen Regressionsanalyse (Einzelzufriedenheiten als unabhängigen Variablen,

⁸⁵ Bei den Qualitätsmerkmalen handelte es sich um Anzahl der Fahrten tagsüber, Anzahl der Fahrten am Wochenende und abends, Sicherheit der Anschlüsse, Pünktlichkeit, Schnelligkeit, Übersichtlichkeit des Netzes, Platzangebot im Fahrzeug, Informationen an Haltestellen, Informationen im Fahrzeug, Preis-Leistungsverhältnis und persönliche Sicherheit an Haltestellen.

⁸⁶ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 59–70 und *Bäumer/Pfeiffer* (2011), S. 46–47.

Globalzufriedenheit als abhängige Variable) indirekt ermittelt. Ein Vergleich der direkt erhobenen und der indirekt ermittelten Wichtigkeiten zeigte starke Unterschiede. Aus Gründen der Plausibilität der Ergebnisse wurde die direkte Abfrage vorgezogen.⁸⁷ Eine Penalty-Reward-Faktoren-Analyse⁸⁸ wurde durchgeführt, bei der hohe Zufriedenheiten (Werte 1 und 2) und niedrigere Zufriedenheiten (Werte 4 und 5) dummykodiert wurden. Es wurden Penalty-Faktoren, aber keine eindeutigen Reward-Faktoren festgestellt.⁸⁹

Mehrere Analysen wurden durchgeführt, um mögliche Wirkungen der angebotsverändernden Maßnahmen zu identifizieren. Dazu wurden die Zufriedenheiten der Vorher- und der Nachher-Erhebung miteinander verglichen. Zwar hat sich die Globalzufriedenheit bei der Mehrzahl der Teilstichproben nicht verändert, aber für manche Teilstichproben konnten signifikante positive ebenso wie signifikante negative Veränderungen festgestellt werden. Diese Tendenz zeigte sich auch hinsichtlich der Einzelzufriedenheiten. Insgesamt konnten daher zumindest teilweise Reaktionen im Zusammenhang mit den qualitätsverändernden Maßnahmen festgestellt werden.⁹⁰

Ein Vergleich der Globalzufriedenheit der Befragten mit veränderter Gruppenzugehörigkeit deutete darauf hin, dass ein Zusammenhang zwischen der Globalzufriedenheit und dem Nutzungsverhalten des ÖPNV besteht.⁹¹ Eine genauere Analyse der von den Befragten genannten Gründe für das veränderte Nutzungsverhalten zeigte jedoch, dass insbesondere Änderungen der persönlichen Lebensumstände wie eine veränderte Pkw-Verfügbarkeit, Umzug oder Wechsel des Arbeitsplatzes

⁸⁷ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 88–91 und *Bäumer/Pfeiffer* (2011), S. 51–52.

⁸⁸ Bei dieser Analyse wird zwischen sogenannten Penalty- und Reward-Faktoren unterschieden. Als Penalty-Faktoren werden Qualitätsmerkmale bezeichnet, deren Abwesenheit zur Unzufriedenheit führt, aber deren Anwesenheit keine Zufriedenheit verursacht. Dahingegen werden als Reward-Faktoren Qualitätsmerkmale bezeichnet, deren Anwesenheit zur Zufriedenheit führt, deren Abwesenheit aber keine Unzufriedenheit hervorruft, vgl. *Homburg/Klarmann* (2006), S. 244–250 und *Matzler/Sauerwein/Stark* (2002), S. 276–278, *Kaiser* (2005), S. 152–155.

⁸⁹ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 92–94 und *Bäumer/Pfeiffer* (2011), S. 53–56.

⁹⁰ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 103–113.

⁹¹ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 117 und *Bäumer/Pfeiffer/Meier* (2012), S. 8–9.

diese Änderungen begründeten, aber auch die durchgeführten Maßnahmen.⁹²

Die Analyse des Zusammenhangs zwischen der Kundenzufriedenheit und dem wirtschaftlichen Erfolg basierte nicht auf den Zufriedenheitsdaten der Vorher-/Nachher-Erhebung. Stattdessen wurden einerseits Kundenzufriedenheitsdaten des ÖPNV-Kundenbarometers in Form von Mittelwerten der Globalzufriedenheit und andererseits vom Verband Deutscher Verkehrsunternehmen (VDV) erhobene Unternehmenskennziffern in einer Zeitreihenanalyse verwendet. Es wurden drei verschiedene Modelle geschätzt, die den Ertrag, die Anzahl der Fahrgäste und die Anzahl der Personenkilometer (als jeweils abhängige Variable) erklären sollten. Als unabhängige Variablen dienten der Mittelwert der Globalzufriedenheit aus dem ÖPNV-Kundenbarometer, die Anzahl der Einwohner im Bedienungsgebiet, die Anzahl der Platzkilometer⁹³ (zur Berücksichtigung von unternehmens- und verkehrsangebotsspezifischen Unterschieden) und das Erhebungsjahr (zur Berücksichtigung eines zeitlichen Trends).⁹⁴

Mit den Ergebnissen der Analysen konnten folgende Zusammenhänge erklärt werden:

- Zwischen dem Erhebungsjahr und dem Ertrag sowie zwischen dem Erhebungsjahr und der Anzahl der Fahrgäste besteht jeweils ein positiver und signifikanter Zusammenhang, d. h. durchschnittlich sind im Zeitverlauf der Ertrag und die Anzahl der Fahrgäste gestiegen. Dahingegen konnte kein signifikanter Zusammenhang zwischen dem Erhebungsjahr und der Anzahl der Personenkilometer festgestellt werden.
- Zwischen der Anzahl der Platzkilometer und allen drei Kennzahlen besteht ein positiver und signifikanter Zusammenhang, d. h. es liegt ein höherer Unternehmenserfolg bei einer größeren Anzahl an Platzkilometern vor.
- Zwischen der Anzahl der Einwohner und den drei Kennzahlen besteht fast durchgehend kein signifikanter Zusammenhang.

⁹² Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 133.

⁹³ Die Platzkilometer beziehen sich auf Sitz- und Stehplätze, vgl. *Verband Deutscher Verkehrsunternehmen* (2012), S. 86.

⁹⁴ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 136–138 und *Bäumer/Pfeiffer/Meier* (2012), S. 11–12.

- Zwischen der mittleren Globalzufriedenheit und dem Ertrag sowie zwischen der mittleren Globalzufriedenheit und der Anzahl der Fahrgäste besteht ein negativer und signifikanter Zusammenhang, d. h. eine höhere Zufriedenheit geht mit einem höheren Ertrag und einer höheren Anzahl Fahrgästen einher. Dahingegen besteht ein solcher Zusammenhang für die Anzahl der Personenkilometer nicht.

Somit konnte zumindest teilweise ein Zusammenhang zwischen Kundenzufriedenheit und wirtschaftlichem Erfolg festgestellt werden. Allerdings wurde auf mögliche Selektionseffekte bedingt durch die freiwillige Teilnahme der Unternehmen bzw. Verkehrsverbände verwiesen.⁹⁵

Diese Analysen leisten einen Beitrag für die empirische Evidenz des Zusammenhangs zwischen Kundenzufriedenheit und wirtschaftlichem Erfolg. Aufgrund der Verwendung der Globalzufriedenheit bleibt der analysierte Zusammenhang jedoch eher allgemein, sodass sich direkt keine Handlungsempfehlungen bezüglich der Veränderung einzelner Qualitätsmerkmale ableiten lassen. Zudem könnten Veränderungen der Kennziffern des wirtschaftlichen Erfolgs auch auf andere Ursachen zurückzuführen sein, die bei der Analyse anhand der Zeitreihen unberücksichtigt blieben, beispielsweise Ausgaben für Marketingaktivitäten und die Verfügbarkeit von anderen Verkehrsmitteln.

2.2.4 Bildung eines Qualitätsindex

Von *Hensher/Stopher/Bullock* (2003) wurde die Bildung eines Qualitätsindex (*Service Quality Index*) vorgeschlagen. Zu diesem Zweck erfolgte eine schriftliche Befragung von Busnutzern unterschiedlicher Busunternehmen und Buslinien. Die Befragung bestand im Wesentlichen aus zwei Teilen: Gegenstand des ersten Teils war die Beschreibung der jeweils aktuell durchgeführten Fahrt in Bezug auf 13 Qualitätsattribute. Im zweiten Teil wurden Verkehrsmittelwahlexperimente (Discrete Choice Experimente⁹⁶) durchgeführt. In einem solchen Experiment standen als Wahlalternativen zwei hypothetische Busalternativen sowie das Busangebot der aktuellen Fahrt zur Verfügung. Die hypothetischen Bus-

⁹⁵ Vgl. *Bäumer/Hussinger/Pfeiffer* (2010), S. 136–140 und *Bäumer/Pfeiffer/Meier* (2012), S. 12.

⁹⁶ Auf Verfahren der Discrete Choice Analyse wird in Kap. 3 eingegangen.

alternativen wurden durch Kombinationen von Ausprägungen der 13 Qualitätsattribute charakterisiert. Zur Auswertung der Daten der Wahlexperimente wurde ein künstliches Nested Logit-Modell verwendet, sodass je nach Erhebungsort Unterschiede in den Skalierungsparametern berücksichtigt werden konnten.⁹⁷ Für die einzelnen Erhebungsorte wurden spezifische Nutzenparameter geschätzt.

Ein Qualitätsindex wurde für jeden einzelnen Befragten ermittelt, indem die Ausprägungen der jeweils real durchgeführten Fahrt mit den Nutzenparametern gewichtet und anschließend aufsummiert wurden. Darauf aufbauend wurden durch Bildung von Mittelwerten Qualitätsindizes für die einzelnen Buslinien bestimmt. Anhand dieser Indizes sollte ein Vergleich der Qualität der unterschiedlichen Buslinien ermöglicht werden.

Zudem wurden Qualitätsattribute identifiziert, die den Qualitätsindex wertmäßig am stärksten nach oben oder unten verändern: Dabei handelt es sich einerseits um die Frequenz und die Sitzplatzverfügbarkeit und andererseits um die Gesamtreisezeit und den Fahrpreis. Die auf diese Weise ermittelten Ergebnisse sollten zur Ableitung von Handlungsempfehlungen dienen.⁹⁸

Dabei ist jedoch zu beachten, dass bei diesen Wahlexperimenten unterschiedliche Busangebote als Alternativen zur Verfügung standen, andere Verkehrsmittel hingegen nicht berücksichtigt wurden. Daher sind Aussagen bezüglich der intramodalen Verkehrsmittelwahl möglich, d. h. Marktanteile von vergleichbaren Buslinien, die sich hinsichtlich der Qualität unterscheiden, könnten bestimmt werden. Zudem lassen sich anhand der Nutzenparameter Zahlungsbereitschaften für die Veränderung der Qualität bestimmen. Da die Wahlexperimente jedoch nicht verkehrsmittelübergreifend gestaltet wurden, sind keine Aussagen bezüglich einer Veränderung des Modal Splits möglich, sodass die Ableitung von Handlungsempfehlungen deutlich eingeschränkt ist.

⁹⁷ Auf ein derartiges Modell und die Bedeutung der Skalierungsparameter wird in Kap. 3.3 und Kap. 3.5.2 eingegangen.

⁹⁸ Vgl. *Hensher/Stopher/Bullock* (2003), S. 501–516. Für eine weitere Anwendung des Qualitätsindex vgl. *Hensher/Prioni* (2002).

2.2.5 Kundenzufriedenheit, Qualität und Verkehrsmittelwahl

Eine derartige Einschränkung besteht bei der von *Angermüller* (2007) durchgeführten Analyse des Zusammenhangs zwischen der Kundenzufriedenheit bzw. der Qualität und der Verkehrsmittelwahl hingegen nicht. Bei dieser Analyse wurden zwei Ansätze verwendet.

Der erste Ansatz basierte auf Wegeprotokollen, in denen Befragte nicht nur ihre real durchgeführten Wege und das dabei verwendete Verkehrsmittel dokumentierten, sondern diese Wege auch unter Verwendung möglicher Verkehrsmittelalternativen beschrieben. Zudem wurde die Zufriedenheit in Bezug auf 26 Merkmale erfragt. Diese Qualitätsmerkmale fungierten neben Eigenschaften der Fahrt und soziodemografischen Charakteristika der Befragten als Attribute in einem Logit-Modell, um den Einfluss auf die Verkehrsmittelwahl zu bestimmen. Es wurde festgestellt, dass die Fahrzeit den stärksten Einfluss auf die Verkehrsmittelwahl ausübt, gefolgt von der Zufriedenheit in Bezug auf Haltestelle, Pünktlichkeit, Informationssystem, Umsteigehäufigkeit und Anschlusssicherheit.

Der zweite Ansatz basierte auf Verkehrsmittelwahlexperimenten (Discrete Choice Experimente), die sich auf eine konkrete Fahrt des Befragten bezogen. Als Wahlalternativen standen zwei öffentliche Verkehrsmittel (allgemein als „Bus oder Bahn“ bezeichnet), die durch Zeit und Preis sowie einzelne Qualitätsattribute⁹⁹ charakterisiert wurden, und eine dritte Alternative (MIV, Fahrrad oder zu Fuß), die nicht weiter charakterisiert wurde. Die Daten wurden mit gängigen Verfahren der Discrete Choice Analyse ausgewertet. Auf Basis der Modellergebnisse wurde ein Simulationsmodell entwickelt, anhand dessen in Abhängigkeit der Qualitätsattribute die Verkehrsmittelwahl bestimmt und der Modal Split sowie die Erlössituation veranschaulicht werden konnten. Als Einsatzmöglichkeit dieses Simulationsmodells wurde die Möglichkeit genannt, unter Berücksichtigung der Kosten eine Priorisierung von Maßnahmen zur Veränderung der Qualität vorzunehmen.¹⁰⁰

⁹⁹ In einem Beispiel, in dem ein solches Wahlexperiment dargestellt wurde, wurden lediglich zwei Qualitätsattribute verwendet, vgl. *Angermüller* (2007), S. 58.

¹⁰⁰ Vgl. *Angermüller* (2007), S. 59–61.

Im Gegensatz zu gängigen Verfahren der Messung der Qualität bzw. der Kundenzufriedenheit berücksichtigen diese beiden Ansätze den Zusammenhang zur Verkehrsmittelwahl explizit. Jedoch besteht beim ersten Ansatz eine Einschränkung dergestalt, dass sich die Analyse aufgrund der Verwendung von Realdaten auf Verkehrsmittel und Qualitätsattribute beschränkt, die in der Realität vorzufinden sind. Zwar lassen sich beim zweiten Ansatz auch (noch) nicht bestehende Verkehrsmittel und Qualitätsattribute berücksichtigen, doch besteht dort eine Einschränkung hinsichtlich der Anzahl der analysierten Qualitätsattribute.

2.3 Zusammenfassung

Verkehrsdienstleistungen im Bereich des öffentlichen Personennahverkehrs, die als Dienstleistungen von allgemeinem (wirtschaftlichem) Interesse bzw. als Aufgabe der Daseinsvorsorge angesehen werden, können vielfach nicht wirtschaftlich betrieben werden, sodass am Markt kein ausreichendes Angebot entstehen würde. Daher werden sie von der öffentlichen Hand bereitgestellt.

Gesetze auf europäischer, nationaler und regionaler Ebene bilden den institutionellen Rahmen des Schienenpersonennahverkehrs. Aufgrund eines fehlenden wettbewerblichen Angebots sind gemeinwirtschaftliche Verkehrsdienstleistungen zu vergeben. Infolgedessen wird in der Regel ein öffentlicher Dienstleistungsauftrag erforderlich, in dem neben Art, Umfang und Dauer der Verkehrsdienstleistungen auch die Aufteilung der Kosten und Einnahmen geregelt werden und Qualitätskriterien berücksichtigt werden können. In Deutschland wird die Erbringung gemeinwirtschaftlicher Verkehrsleistungen an die Länder übertragen, wobei in Nordrhein-Westfalen wiederum die Planung, Organisation und Ausgestaltung des SPNV an die Kreise und kreisfreien Städte übertragen wird, die sich zu einem Zweckverband oder einer gemeinsamen Anstalt zusammenschließen. Es werden in NRW auch lenkungspolitische Ziele dergestalt verfolgt, dass die Qualität und Quantität des Verkehrsangebots auf der Schiene verbessert und dadurch das Verkehrsaufkommen des SPNV gesteigert werden soll.

Im Nahverkehrsplan des Zweckverbands Nahverkehr Westfalen-Lippe werden Grundstandards und anzustrebende Ziele hinsichtlich der Qualität und Quantität des Verkehrsangebots festgelegt, die als Grundlage für Wettbewerbsverfahren dienen. Eine Überprüfung der erbrachten Leistung mit geeigneten Messverfahren, u. a. Kundenzufriedenheitsmessungen, ist vorgesehen.

Grundsätzlich können für die Gestaltung eines Verkehrsvertrags unterschiedliche Vertragsformen (Bruttovertrag, Nettovertrag und Mischvertrag) eingesetzt werden, die sich im Wesentlichen dahingehend unterscheiden, wie das Einnahmerisiko aus Fahrgeldeinnahmen zwischen dem Aufgabenträger und dem Verkehrsunternehmen verteilt ist, sodass unterschiedliche Anreizstrukturen bestehen. Je nach verwendeter Vertragsform ist die Kenntnis des Zusammenhangs zwischen der Qualität und der Verkehrsmittelwahl für den Aufgabenträger und/oder für das Verkehrsunternehmen und damit aus gemeinwirtschaftlicher und/oder einzelwirtschaftlicher Sicht von Bedeutung.

Bei gängigen Verfahren zur Messung der Kundenzufriedenheit wird die relative Wichtigkeit der Qualitätsmerkmale entweder direkt erfragt oder indirekt ermittelt. Sogenannte Kundenzufriedenheitsportfolios sollen dabei helfen, in Abhängigkeit der relativen Wichtigkeit und der Kundenzufriedenheit für einzelne Qualitätsmerkmale Handlungsempfehlungen abzuleiten. Dabei muss jedoch unterstellt werden, dass ein Zusammenhang zwischen der Kundenzufriedenheit und der Verkehrsmittelwahl bzw. der Verkehrsmittelnutzung besteht. Dieser Zusammenhang ist bislang empirisch nur unzureichend belegt.

Exemplarisch wurden drei sehr unterschiedliche Studien mit einem Bezug zur Verkehrsmittelwahl bzw. zum wirtschaftlichen Erfolg vorgestellt. In der ersten Studie wurde einerseits der Zusammenhang zwischen der Kundenzufriedenheit hinsichtlich einzelner Qualitätsmerkmale und der Globalzufriedenheit und andererseits der Zusammenhang zwischen der Globalzufriedenheit und dem wirtschaftlichen Erfolg analysiert. Ein Zusammenhang zum wirtschaftlichen Erfolg konnte zumindest teilweise hergestellt werden; dieser blieb jedoch sehr allgemein. In der zweiten Studie wurden Verkehrsmittelwahlexperimente durchgeführt und aufbauend auf den Modellergebnissen und den Merkmalen von Fahrten

Qualitätsindizes gebildet, anhand derer Handlungsempfehlungen abgeleitet werden sollten. Allerdings ist die Ableitung von Handlungsempfehlungen dabei deutlich eingeschränkt, da keine verkehrsmittelübergreifenden Wahlexperimente verwendet wurden. In der dritten Studie basierte die Analyse der Verkehrsmittelwahl zum einen auf Wegeprotokollen, zum anderen auf Verkehrsmittelwahlexperimenten, dessen Modellergebnisse in einer Simulation zur Ableitung von Handlungsempfehlungen verwendet wurden. Bei ersterem besteht eine Einschränkung hinsichtlich der real existierenden Verkehrsmittel und Qualitätsattribute und bei letzterem hinsichtlich der Anzahl der verwendeten Qualitätsattribute.

3 Methodische Grundlagen der Discrete Choice Analyse

3.1 Allgemeine Grundlagen der Discrete Choice Analyse

Die Verkehrsmittelwahl wird vielfach unter Verwendung der Discrete Choice Analyse (teilweise auch als auswahlbasierte oder Choice Based Conjoint Analyse bezeichnet)¹⁰¹ untersucht. Die Discrete Choice Analyse gehört zu den dekompositionellen Verfahren zur Analyse von Nutzenstrukturen von Individuen. Im Gegensatz zu den kompositionellen Verfahren, bei denen Bewertungen von Eigenschaften eines Produktes (oder eines Verkehrsmittels) getrennt erhoben und anschließend zu einem Gesamturteil verknüpft werden, erfolgt bei den dekompositionellen Verfahren die Erhebung eines Gesamturteils, aus dem auf die Bedeutung der einzelnen Eigenschaften (Attribute) geschlossen wird.¹⁰²

Bei der Discrete Choice Analyse handelt es sich bei den erhobenen Gesamturteilen um diskrete Wahlentscheidungen, d. h. ein Individuum n wählt genau eine von mehreren verfügbaren Alternativen.¹⁰³ Es wird unterstellt, dass das Individuum nutzenmaximierend agiert und somit die Alternative wählt, die ihm den größten Nutzen stiftet.¹⁰⁴

Im Gegensatz zur traditionellen Conjoint-Analyse, welche auf einem deterministischem Nutzenkonzept beruht und damit unterstellt, dass alle bestehenden Einflüsse auf die Bewertung explizit im Modell abgebildet werden, beruht die Discrete Choice Analyse auf dem Zufallsnutzenkonzept (*Random Utility Theory*). Demnach besteht der Nutzen U_{jn} , den die Alternative j dem Individuum n stiftet, aus einer deterministischen

¹⁰¹ In der englischsprachigen Literatur wird versucht, durch Verwendung der Bezeichnung „Discrete Choice Analysis“ eine deutliche Abgrenzung von der Conjoint-Analyse vorzunehmen, vgl. *Louviere/Flynn/Carson* (2010), S. 61–62.

¹⁰² Vgl. *Helm/Steiner* (2008), S. 213–216 und *Backhaus/Erichson/Weiber* (2011), S. 318.

¹⁰³ Dabei kann es sich bei einer dieser Alternativen auch um die sogenannte „no-option“ handeln, wodurch ausgedrückt wird, dass keine der anderen Alternativen gewählt wird.

¹⁰⁴ Vgl. *Train* (2009), S. 14.

(systematischen) Komponente V_{jn} und einer stochastischen (zufälligen) Komponenten ε_{jn} .¹⁰⁵

$$U_{jn} = V_{jn} + \varepsilon_{jn}. \quad (3.1)$$

Die Verwendung einer stochastischen Komponente bedeutet nicht, dass sich die Individuen bei ihren Entscheidungen unsicher sind oder die Entscheidungen gar zufällig treffen; sie können sehr konsistente und systematische Präferenzen besitzen und ihren Nutzen vollkommen deterministisch maximieren. Stattdessen wird durch die Verwendung der stochastischen Nutzenkomponente ausgedrückt, dass – im Gegensatz zu traditionellen ökonomischen Modellen – nicht der Anspruch erhoben wird, dass alle bestehenden Einflüsse auf die Wahlentscheidung im Modell explizit abgebildet werden können.¹⁰⁶ Die nicht beobachteten Einflüsse werden durch die stochastische Komponente berücksichtigt. Diese komplexere, aber realistischere Annahme ist mit einem höheren Aufwand bei der Modellierung der Wahlentscheidungen verbunden.

Die Beschreibung der systematischen Nutzenkomponente V_{jn} erfolgt unter Verwendung des konsumtheoretischen Ansatzes nach *Lancaster* (1966). Entsprechend wird angenommen, dass nicht die Alternative j an sich, sondern deren Eigenschaften (Attributsausprägungen) Nutzen stiften.¹⁰⁷ Die systematische Nutzenkomponente V_{jn} wird mathematisch in einer Nutzenfunktion abgebildet. Mit dieser Nutzenfunktion wird ausgedrückt, wie die einzelnen Teilnutzen, die mit den Attributsausprägungen verbunden sind, miteinander verknüpft werden. Im Rahmen der Modellierung gilt es, geeignete Teilnutzenmodelle auszuwählen, die additiv oder multiplikativ miteinander verknüpft werden.

Es können grundlegend drei Teilnutzenmodelle unterschieden werden. Dabei handelt es sich um das Vektor-, das Idealpunkt- und das Teilwert-Modell (Parthworth-Modell).¹⁰⁸ Beim Vektor-Modell wird ein positiver (oder negativer) linearer Zusammenhang zwischen der Attributsausprägung und dessen Nutzen unterstellt, sodass mit jeder weiteren

¹⁰⁵ Vgl. *Hensher/Rose/Greene* (2005), S. 75 und *Louviere et al.* (2000), S. 38.

¹⁰⁶ Vgl. *Louviere et al.* (2000), S. 39 und *Train* (2009), S. 14–15.

¹⁰⁷ Vgl. *Lancaster* (1966), S. 133 und *Ben-Akiva/Lerman* (1987), S. 47–48.

¹⁰⁸ Vgl. *Green* (1978), S. 105–107 und *Backhaus/Erichson/Weiber* (2011), S. 329–330.

Erhöhung (oder Absenkung) der Nutzen weiter steigt (oder sinkt). Dahingegen unterstellt das Idealpunkt-Modell, dass eine ideale Attributsausprägung existiert, mit dem der größte Nutzen verbunden ist; eine Über- oder Unterschreitung stiftet einen geringeren Nutzen. Bei diesen beiden Teilnutzenmodellen besteht ein stetiger Nutzenverlauf. Dahingegen werden beim Teilwert-Modell den diskreten Ausprägungen eines Attributs diskrete Nutzenwerte zugeordnet. Das Vektor- und das Idealpunktmodell können auch als Spezialfälle des Teilwert-Modells interpretiert werden.¹⁰⁹

Üblicherweise wird zur Beschreibung der systematischen Nutzenkomponente eine linear additive Nutzenfunktion der Form

$$V_{jn} = \sum_{m=1}^M \sum_{l=1}^{L-1} \beta_{jlmn} x_{jlmn} \quad (3.2)$$

verwendet, wobei es sich bei x_{jlmn} um die (entsprechend kodierte) l -te Ausprägung der 1 bis M Attribute und bei β_{jlmn} um den zugehörigen Nutzenparameter handelt. Beim Teilwert-Modell wird jede (kodierte) Ausprägung¹¹⁰ mit einem eigenen Nutzenparameter, beim Vektormodell hingegen jede Ausprägung eines Attributs mit dem gleichen Nutzenparameter verknüpft.

Die Nutzenfunktion in (3.2) lässt sich übersichtlicher in Vektorschreibweise ausdrücken:

$$V_{jn} = \boldsymbol{\beta}'_{jn} \mathbf{X}_{jn}, \quad (3.3)$$

wobei es sich bei $\boldsymbol{\beta}_{jn}$ um einen Spaltenvektor, der die Nutzenparameter beinhaltet, und bei \mathbf{X}_{jn} um einen Spaltenvektor, der die Ausprägungen der Attribute beinhaltet, handelt.¹¹¹

Mit ‚linear‘ ist bei der Nutzenfunktion gemeint, dass sie linear in den Parametern (in diesem Fall in den Nutzenparametern) ist. Die Attributsausprägungen sind hingegen nicht notwendigerweise linear. Anstelle von linearen Attributsausprägungen können mitunter auch die Logarithmen

¹⁰⁹ Vgl. *Green* (1978), S. 106.

¹¹⁰ Hierbei ist zu beachten, dass zur Kodierung eines qualitativen Attributs mit insgesamt L Ausprägungen insgesamt $L-1$ Kodierungsvariablen verwendet werden.

¹¹¹ Vgl. *Louviere et al.* (2000), S. 49. Mit dem Apostroph wird die Transponierung des Vektors gekennzeichnet.

oder Potenzen der Ausprägungen verwendet werden.¹¹² Verbunden mit der Verwendung einer additiven Verknüpfung ist die Annahme, dass ein geringerer Nutzenbetrag eines Attributs durch einen größeren Nutzenbetrag eines anderen Attributs kompensiert werden kann (kompensatorisches Nutzenmodell).¹¹³

Bei den erhobenen Wahlentscheidungen kann es sich entweder um Realdaten (*Revealed Preference*) oder um Experimentaldaten (*Stated Preference/Stated Choice*) handeln.¹¹⁴ Realdaten sind insbesondere durch eine hohe Reliabilität und Augenscheinvalidität (*Face Validity*)¹¹⁵ gekennzeichnet. Allerdings sind die Analysemöglichkeiten mit dieser Art von Daten dahingehend eingeschränkt, dass nur derzeit bzw. in der Vergangenheit existierende Alternativen mit existierenden Attributen und Attributsausprägungen analysiert werden können. Zudem sind bei Realdaten die Ausprägungen der Attribute oftmals miteinander korreliert. So sind beispielsweise im Schienenpersonenregionalverkehr eine niedrigere Qualität mit einem niedrigeren Preis (zweite Klasse) und eine höhere Qualität mit einem höheren Preis (erste Klasse) verbunden. In Experimenten mit fiktiven Wahlsituationen lassen sich jedoch nicht nur bestehende, sondern auch neue Alternativen, Attribute und/oder Attributsausprägungen kontrolliert einbeziehen. Eine realitätsnahe Gestaltung der fiktiven Wahlsituation ist von Bedeutung, da als unrealistisch wahrgenommene Wahlsituationen die Datenqualität negativ beeinflussen.

Durch eine Einbeziehung von Realdaten in Wahlexperimente werden die Vorteile der Realdaten mit den Vorteilen der Experimentaldaten kombiniert. Auf diese Weise ist in den Experimenten ein Bezug zu einer tatsächlichen Handlung möglich. In einem Verkehrsmittelwahlexperiment lässt sich ein Bezug zu einer konkreten Fahrt herstellen, welche zunächst beschrieben wird und deren Ausprägungen anschließend in realistischen

¹¹² Vgl. *Louviere et al.* (2000), S. 42.

¹¹³ Im Gegensatz dazu ist bei den nicht-kompensatorischen Nutzenmodellen, zu denen das konjunktive, das subjunktive und das lexikografische Verknüpfungsmodell zählen, eine derartige Kompensation nicht möglich. Beispielsweise kann das nicht Vorhandensein einer bestimmten Ausprägung zur Ablehnung der Alternative führen, vgl. *Fischer* (2001), S. 69–72.

¹¹⁴ Vgl. *Hensher/Rose/Greene* (2005), S. 92–97.

¹¹⁵ Vgl. *Moosbrugger/Kelava* (2007), S. 11–15.

Schwankungsbreiten variiert werden.¹¹⁶ Derartige Wahlsituationen sind in der Regel leichter vorstellbar als rein hypothetische Wahlsituationen.

3.2 Binomiales und multinomiales Logit-Modell

Im einfachsten Fall erfolgt eine Wahl zwischen zwei Alternativen. Diese beiden Alternativen stellen das Choice Set C_n dar und werden mit i und j bezeichnet. Aus der Annahme, dass sich ein Individuum n zufallsnutzenmaximierend verhält, folgt, dass es die Alternative i genau dann wählt, wenn diese einen größeren Zufallsnutzen aufweist als Alternative j . Aufgrund der Verwendung des Zufallsnutzen kann die Wahl einer Alternative jedoch nicht vollständig, sondern nur zu einer bestimmten Wahrscheinlichkeit angegeben werden. Dabei ist die Wahrscheinlichkeit, dass Alternative i gewählt wird, gleich der Wahrscheinlichkeit, dass Alternative i einen größeren Nutzen als Alternative j stiftet.¹¹⁷

$$P_n(i | C_n) = P_n(i) = P(U_{in} \geq U_{jn}). \quad (3.4)$$

Die Wahrscheinlichkeit, dass Alternative j gewählt wird, beträgt entsprechend

$$P_n(j) = 1 - P_n(i). \quad (3.5)$$

Wird Gleichung (3.1) in (3.4) eingesetzt, ergibt sich

$$P_n(i) = (V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}). \quad (3.6)$$

Die Gleichung lässt sich umformen zu

$$P_n(i) = (\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}). \quad (3.7)$$

Dies bedeutet, dass die Wahrscheinlichkeit, dass Alternative i gewählt wird, gleich der Wahrscheinlichkeit ist, dass die Differenz der zufälligen Nutzenkomponenten $(\varepsilon_{jn} - \varepsilon_{in})$ kleiner ist als die Differenz der determinis-

¹¹⁶ Vgl. *Hensher* (2004), S. 429–431, *Caussade et al.* (2005), S. 627 und *Vrtic/Fröhlich* (2006), S. 52–53.

¹¹⁷ Vgl. *Louviere et al.* (2000), S. 40.

tischen Nutzenkomponenten $(V_{in} - V_{jn})$.¹¹⁸ Da die zufälligen Nutzenkomponenten sowie deren Differenz nicht beobachtet werden können, muss eine Annahme über die Verteilung dieser Zufallsvariablen getroffen werden. Je nach angenommener Verteilungsfunktion der stochastischen Nutzenkomponenten ergeben sich unterschiedliche Modelltypen.

Unter der Annahme, dass ε_{jn} und ε_{in} unabhängig voneinander verteilt, identisch verteilt und Gumbel-verteilt¹¹⁹ sind, folgt, dass deren Differenz $\varepsilon_n = (\varepsilon_{jn} - \varepsilon_{in})$ logistisch¹²⁰ verteilt ist mit der Dichtefunktion

$$f(\varepsilon_n) = \frac{\mu e^{-\mu\varepsilon_n}}{(1+e^{-\mu\varepsilon_n})^2} \quad (3.8)$$

und der Verteilungsfunktion

$$F(\varepsilon_n) = \frac{1}{1+e^{-\mu\varepsilon_n}}, \quad \mu > 0, \quad -\infty < \varepsilon_n < \infty, \quad (3.9)$$

wobei es sich bei μ um einen positiven Skalierungsparameter handelt. Entsprechend beträgt die Wahrscheinlichkeit, dass Alternative i gewählt wird¹²¹

$$P_n(i) = P(U_{in} \geq U_{jn}) = \frac{1}{1+e^{-\mu(V_{in}-V_{jn})}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}} + e^{\mu V_{jn}}}. \quad (3.10)$$

Dieses Modell wird als binomiales Logit-Modell bezeichnet.

In Analogie zum binomialen Logit-Modell lässt sich das multinomiale Logit-Modell mit mehr als zwei Alternativen herleiten. Alternative i wird dabei mit einer sogenannten „zusammengesetzten“ Alternative (*composite alternative*) verglichen. Diese „zusammengesetzte“ Alternative umfasst alle anderen Alternativen des Choice Sets C_n und weist den Zufallsnutzen der besten von diesen Alternativen auf.¹²² Alternative i wird gewählt,

¹¹⁸ Vgl. Ben-Akiva/Lerman (1987), S. 61.

¹¹⁹ Die Gumbel-Verteilung wird auch als Extremwertverteilung-Typ-1 bezeichnet.

¹²⁰ Die logistische Verteilung ist der Normalverteilung sehr ähnlich, wobei die logistische Verteilung etwas breitere Verteilungsenden hat (*fatter tails*), vgl. Ben-Akiva/Lerman (1987), S. 71.

¹²¹ Vgl. Ben-Akiva/Lerman (1987), S. 71.

¹²² Vgl. Ben-Akiva/Lerman (1987), S. 103.

wenn sie einen größeren Nutzen als die „zusammengesetzte“ Alternative stiftet:

$$P_n(i | C_n) = P_n(i) = P(U_{in} \geq \max_{\substack{j \in C_n \\ j \neq i}} U_{jn}). \quad (3.11)$$

Die Wahlwahrscheinlichkeit der Alternative i beträgt

$$P_n(i) = P \left[V_{in} + \varepsilon_{in} \geq \max_{\substack{j \in C_n \\ j \neq i}} (V_{jn} + \varepsilon_{jn}) \right]. \quad (3.12)$$

Es wird angenommen, dass ε_{in} und alle ε_{jn} unabhängig verteilt, identisch verteilt und Gumbel-verteilt sind. Aufgrund der Eigenschaften, dass eine Gumbel-Verteilung auch bei linearen Transformationen bestehen bleibt, das Maximum von unabhängig und Gumbel-verteilten Zufallsvariablen ebenfalls Gumbel-verteilt ist, und die Differenz zweier unabhängig und Gumbel-verteilten Zufallsvariablen logistisch verteilt ist, ergibt sich das multinomiale Logit-Modell.¹²³ Die Wahrscheinlichkeit, dass Alternative i gewählt wird, beträgt

$$P_n(i) = \frac{1}{\sum_{j=1}^J e^{-\mu(V_{in}-V_{jn})}} = \frac{e^{\mu V_{in}}}{\sum_{j=1}^J e^{\mu V_{jn}}}, \quad (3.13)$$

wobei es sich bei J um die Anzahl der Alternativen im Choice Set C_n und bei μ um einen positiven Skalierungsparameter handelt. Hieraus wird ersichtlich, dass das binomiale Logit-Modell einen Spezialfall des multinomialen Logit-Modells mit zwei Alternativen darstellt. Die folgenden Ausführungen beziehen sich daher gleichermaßen auf beide Modelle.

Aus der Gleichung (3.13) wird deutlich, dass die Wahlwahrscheinlichkeit einer Alternative maßgeblich von der Differenz der systematischen Nutzenkomponenten der Alternativen abhängt. Durch Einsetzen der linear additiven Nutzenfunktionen aus Gleichung (3.3) wird die Wahlwahrscheinlichkeit der Alternative i beschrieben durch¹²⁴

¹²³ Vgl. *Ben-Akiva/Lerman* (1987), S. 105–108.

¹²⁴ Vgl. *Ben-Akiva/Lerman* (1987), S. 108. Die Normalisierung auf 1 bedeutet, dass die Varianz der stochastischen Nutzenkomponenten $\pi^2/6$ beträgt, vgl. dazu *Train* (2009), S. 23–24 und *Ben-Akiva/Lerman* (1987), S. 71.

$$P_n(i) = \frac{e^{\mu \beta' i_n \mathbf{X}_{in}}}{\sum_{j=1}^J e^{\mu \beta' j_n \mathbf{X}_{jn}}}. \quad (3.14)$$

Die im Vektor \mathbf{X}_{jn} enthaltenden Attributsausprägungen der Alternativen sind beobachtbar, die im Vektor β enthaltenden Nutzenparameter gilt es zu schätzen. Der Skalierungsparameter μ lässt sich in einem einzelnen bi- oder multinomialen Logit-Modell nicht identifizieren und wird daher üblicherweise (willkürlich) auf 1 gesetzt.¹²⁵ Entsprechend wird dieser Parameter in der Literatur nicht immer explizit aufgeführt.¹²⁶

Bedingt durch die Eigenschaft der unabhängig und identisch verteilten (*independently and identically distributed*, IID) stochastischen Nutzenkomponenten wird für den Nutzen jeder Alternative der gleiche Skalierungsparameter μ verwendet.¹²⁷ Diese Eigenschaft impliziert, dass die Varianzen der stochastischen Nutzenkomponenten der Alternativen identisch sind und nicht miteinander korrelieren, d. h. in einer Kovarianzmatrix betragen die Werte der Kovarianzen (d. h. alle Nichtdiagonalelemente) 0.¹²⁸ Dadurch ergeben sich rechnerische Vorteile. Die IID-Eigenschaft geht jedoch mit der Eigenschaft der Unabhängigkeit von irrelevanten Alternativen (*Independence of/from irrelevant alternatives*, IIA) einher. Diese besagt, dass das Verhältnis der Wahlwahrscheinlichkeiten von zwei Alternativen unabhängig von dem systematischen Nutzen weiterer Alternativen (bzw. der Abwesenheit oder Anwesenheit dieser Alternativen) ist:¹²⁹

$$\frac{P(i|\tilde{C}_n)}{P(j|\tilde{C}_n)} = \frac{P(i|C_n)}{P(j|C_n)} \quad i, j \in \tilde{C}_n \subseteq C_n. \quad (3.15)$$

Zur Veranschaulichung der Verletzung der IIA-Annahme kann das vielfach zitierte „Red Bus/Blue Bus“-Paradoxon herangezogen werden.¹³⁰ In der Ausgangssituation liegen dabei nur die beiden Alternativen (roter) Bus und Pkw mit einer jeweiligen Wahlwahrscheinlichkeit von $P_n(\text{Pkw}) =$

¹²⁵ Vgl. *Ben-Akiva/Lerman* (1987), S. 107.

¹²⁶ Vgl. beispielsweise *Louviere et al.* (2000), S. 49 und *Ronning* (1991), S. 38.

¹²⁷ Vgl. *Ben-Akiva/Lerman* (1987), S. 104.

¹²⁸ Vgl. *Louviere et al.* (2000), S. 15.

¹²⁹ Vgl. *McFadden* (1974), S. 109 und *Ben-Akiva/Lerman* (1987), S. 108.

¹³⁰ Vgl. dazu beispielsweise *McFadden* (1974), S. 113 und *Ben-Akiva/Lerman* (1987), S. 51–55.

$P_n(\text{roter Bus}) = 0,5$. Das Verhältnis der Wahlwahrscheinlichkeiten dieser beiden Alternativen beträgt somit $P_n(\text{roter Bus})/P_n(\text{Pkw}) = 0,5/0,5 = 1$. Anschließend wird dem Choice Set eine weitere Alternative in Form einer Busalternative hinzugefügt, die sich von der bestehenden Alternative nur anhand der Farbe (blau) unterscheidet. Unter der Annahme, das Verhältnis der Wahlwahrscheinlichkeiten sei konstant und die Wahlwahrscheinlichkeit der beiden Busalternativen gleich groß, muss gelten $P_n(\text{Pkw}) = P_n(\text{roter Bus}) = P_n(\text{blauer Bus}) = 0, \bar{3}$. Insgesamt beträgt die Wahlwahrscheinlichkeit der beiden Busalternativen somit $0, \bar{6}$ und die Wahlwahrscheinlichkeit des Pkw ist vergleichsweise geringer. Es ist realistischer, dass die beiden Busalternativen als eine Alternative wahrgenommen werden, sodass gilt $P_n(\text{Pkw}) = 0,5$ und $P_n(\text{roter Bus}) = P_n(\text{blauer Bus}) = 0,25$. In diesem Fall würde sich jedoch mit $P_n(\text{roter Bus})/P_n(\text{Pkw}) = 0,25/0,5 = 0,5$ das Verhältnis der Wahlwahrscheinlichkeiten ändern, sodass Gleichung (3.15) nicht erfüllt ist.

Bei diesem Beispiel teilen die beiden Busalternativen alle nicht beobachteten Attribute des Busses, sodass die stochastischen Nutzenkomponenten vollständig miteinander korreliert sind. Ein Modell, welches auf der Annahme unabhängig verteilter stochastischer Nutzenkomponenten basiert, würde daher zu verzerrten Nutzenparametern und Fehlern in den Prognosen führen.¹³¹ Während es in diesem Beispiel deutlich erkennbar ist, dass zwischen den Busalternativen Gemeinsamkeiten bestehen und die stochastischen Nutzenkomponenten miteinander korreliert sind, ist dies bei anderen Alternativen weniger deutlich erkennbar. Ob die IIA-Annahme verletzt ist, lässt sich jedoch mit einem entsprechenden Testverfahren überprüfen (vgl. Kap. 3.5.1).

Die IIA-Eigenschaft an sich ist grundsätzlich weder eine erwünschte noch unerwünschte Eigenschaft.¹³² Allerdings ist gegebenenfalls ein entsprechendes Modell zu verwenden, welches nicht auf dieser Eigenschaft beruht.

Das multinomiale Logit-Modell basiert zudem auf der Annahme der Verhaltenshomogenität. Theoretisch ist es möglich, in Experimenten hin-

¹³¹ Vgl. *Ben-Akiva/Lerman* (1987), S. 109 und *Louviere et al.* (2000), S. 45.

¹³² Vgl. *Louviere et al.* (2000), S. 45.

reichend viele Wahlentscheidungen eines Befragten zu erheben, um eine individuelle Nutzenfunktion zu bestimmen. Empirisch ist die Anzahl der abfragbaren Wahlsituation jedoch in der Regel begrenzt, sodass die Bestimmung der Nutzenfunktion aggregiert für alle Befragten oder eine möglichst homogene Teilgruppe¹³³ erfolgt. Ob sich die Nutzenfunktionen von Teilgruppen signifikant unterscheiden, kann mit entsprechenden Testverfahren überprüft werden (vgl. Kap. 3.5.2).

Zur Bestimmung der Nutzenparameter wird üblicherweise das Maximum Likelihood-Schätzverfahren eingesetzt. Dabei gilt es, die unbekannt Parameter so zu bestimmen, dass die Wahlentscheidungen bestmöglich abgebildet werden, d. h. die berechnete Wahlwahrscheinlichkeit der jeweils tatsächlich gewählten Alternative möglichst groß ist.

Unter der Annahme, dass die Wahlentscheidungen voneinander unabhängig sind, beträgt die Likelihood-Funktion, die es zu maximieren gilt

$$L(\boldsymbol{\beta}) = \prod_{n=1}^N \prod_{j=1}^J (P_{nj})^{y_{nj}}, \quad (3.16)$$

wobei es sich bei y_{nj} um eine Indikatorvariable handelt, die 1 beträgt, wenn das Individuum n Alternative j gewählt hat und andernfalls 0.¹³⁴

Anstelle der Likelihood-Funktion L wird üblicherweise die logarithmierte Likelihood-Funktion, bezeichnet als Log-Likelihood-Funktion LL , maximiert.¹³⁵ Die Maximierung der beiden Funktionen führt zu dem gleichen Ergebnis, jedoch ergeben sich bei der Log-Likelihood-Funktion rechnerische Vorteile, da Summen anstelle von Produkten zu maximieren sind. Die Log-Likelihood-Funktion lautet:

$$LL(\boldsymbol{\beta}) = \sum_{n=1}^N \sum_{j=1}^J y_{nj} \ln P_{nj}. \quad (3.17)$$

¹³³ Vgl. *Louviere et al.* (2000), S. 52 und *Ben-Akiva/Lerman* (1987), S. 138.

¹³⁴ In diesem Zusammenhang werden aus Gründen der Übersichtlichkeit mehrere Wahlentscheidungen eines Individuums als jeweils eine Wahlentscheidung unterschiedlicher Individuen abgebildet.

¹³⁵ Vgl. *Ben-Akiva/Lerman* (1987), S. 80–81.

Das Maximum der Log-Likelihood-Funktion¹³⁶ liegt vor, wenn die einzelnen im Vektor β enthaltenen Nutzenparameter β so bestimmt sind, dass die partiellen Ableitungen der Log-Likelihood-Funktion nach den einzelnen Nutzenparametern gleichzeitig 0 betragen:

$$\frac{\partial LL(\beta)}{\partial \beta} \stackrel{!}{=} 0. \quad (3.18)$$

Bei dem so gefundenen relativen Maximum handelt es sich gleichzeitig um ein absolutes Maximum, da die Log-Likelihood-Funktion für Nutzenfunktionen, die linear in den Parametern sind, global konkav ist.¹³⁷

Zur Bestimmung der Nutzenparameter, welche die Log-Likelihood-Funktion maximieren, ist die Verwendung eines iterativen Algorithmus erforderlich. Zu den iterativen Standard-Algorithmen zählen die Newton-Methode und deren Weiterentwicklung/en in Form des (modifizierten) Newton-Raphson-Algorithmus, sowie die BHHH- (Bernd-Hall-Hall-Hausman), DFP- (Davidson-Fletcher-Powell) und BFGS- (Broyden-Fletcher-Goldfarb-Shanno) Algorithmen.¹³⁸

Ausgehend von den Schätzergebnissen für die Nutzenparameter lassen sich Zahlungsbereitschaften bestimmen. In einfachen linearen Modellen handelt es sich dabei um das Verhältnis zweier Nutzenparameter. Sofern eines der zugehörigen Attribute in Geldeinheiten angegeben ist, gibt die Zahlungsbereitschaft an, welchen Wert Individuen für eine bessere Ausprägung der entsprechenden Attributsausprägung zu zahlen bereit sind (*willingness to pay*) bzw. welche Kompensationszahlung für eine schlechtere Ausprägung erforderlich ist (*willingness to accept*):

$$WTP = \frac{\partial U / \partial x_{\text{Qualität}}}{\partial U / \partial x_{\text{Preis}}} = \frac{\beta_{\text{Qualität}}}{\beta_{\text{Preis}}} \quad (3.19)$$

¹³⁶ Die Log-Likelihood-Funktion weist durchgehend negative Werte auf. Diese Eigenschaft resultiert daraus, dass der natürliche Logarithmus im Wertebereich der Wahrscheinlichkeiten ($0 < P_{nj} < 1$) negativ ist.

¹³⁷ Diese Eigenschaft wurde von *McFadden* (1974), S. 115–116 nachgewiesen.

¹³⁸ Vgl. zu diesen Algorithmen *Train* (2009), S. 187–198 und *Gould/Pitblado/Poi* (2010), S. 13–19.

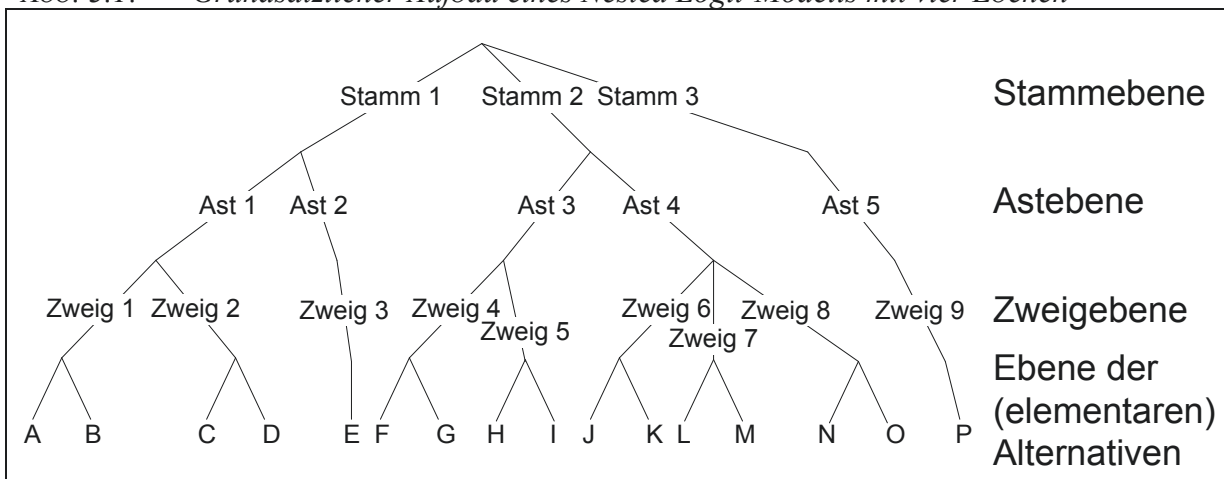
Derartige Veränderungen der Ausprägungen sind nutzenneutral, d. h. der Nutzen verändert sich nicht und somit bleibt auch die Wahlwahrscheinlichkeit gleich.¹³⁹

3.3 Nested Logit-Modell

Im Gegensatz zum multinomialen Logit-Modell erfolgt bei den Modellen der generalisierten Extremwert- (*Generalised Extreme Value*, GEV) Modellfamilie eine teilweise Lockerung der IID-/IIA-Eigenschaft. Die Gemeinsamkeit dieser Modellfamilie besteht darin, dass die stochastischen Nutzenkomponenten gemeinsam als generalisierter Extremwert verteilt sind.¹⁴⁰ Das Nested Logit-Modell stellt einen Spezialfall der GEV-Modellfamilie dar.

Im Nested Logit-Modell werden Wahlalternativen zu sich nicht überlappenden Gruppen, welche als ein Nest bezeichnet werden, zusammengefasst. Der Aufbau des Modells lässt sich als umgekehrte Baumstruktur abbilden (Abb. 3.1) und entspricht einer hierarchischen Verknüpfung von multinomialen Logit-Modellen. Die unterste Ebene wird als Ebene der (elementaren) Alternativen, die Ebenen darüber als Zweig-, Ast- und Stammebene bezeichnet.¹⁴¹

Abb. 3.1: Grundsätzlicher Aufbau eines Nested Logit-Modells mit vier Ebenen



Quelle: Eigene Darstellung in Anlehnung an Hensher/Rose/Greene (2005, S. 496).

¹³⁹ Vgl. Louviere et al. (2000), S. 61 und Hensher/Rose/Greene (2005), S. 357–358.

¹⁴⁰ Vgl. Train (2009), S. 76. Bei der generalisierten Extremwertverteilung handelt es sich um eine Verallgemeinerung der Extremwertverteilungen vom Typ I (Gumbel), Typ II (Fréquet) und Typ III (Weibull), vgl. Coles (2004), S. 46–48.

¹⁴¹ Vgl. Hensher/Rose/Greene (2005), S. 495–496.

Dabei ist jedoch zu beachten, dass es sich dabei nicht um einen Entscheidungsbaum, sondern um ein Wahrscheinlichkeitsmodell zur Abbildung von Varianzunterschieden in den stochastischen Nutzenkomponenten handelt. Das bedeutet, dass bei der Wahl einer Alternative nicht (notwendigerweise) zunächst eine Wahl zwischen den unterschiedlichen Nestern erfolgt und im Anschluss daran eine Wahl zwischen den in diesem Nest enthaltenen Alternativen. Stattdessen wird angenommen, dass die Wahl der elementaren Alternativen simultan zwischen allen Alternativen der unterschiedlichen Nester erfolgt.¹⁴²

Im Folgenden wird ein zweistufiges Nested Logit-Modell beschrieben, dass aus insgesamt W Nestern besteht, die mit $w = 1, \dots, W$ bezeichnet werden. Die Alternativen eines Nests w teilen eine gemeinsame Zufallsnutzenkomponente U_w .¹⁴³ Darüber hinaus weist eine Alternative i des Nests w den Zufallsnutzen $U_{i|w}$ auf. Insgesamt beträgt der Zufallsnutzen der Alternative i im Nest w

$$U_{iw} = U_w + U_{i|w}. \quad (3.20)$$

Die Nutzenkomponenten der beiden Ebenen setzen sich aus den systematischen Nutzenkomponenten V_w und $V_{i|w}$ sowie den stochastischen Nutzenkomponenten ε_w und $\varepsilon_{i|w}$ zusammen

$$U_{iw} = V_w + V_{i|w} + \varepsilon_w + \varepsilon_{i|w}.^{144} \quad (3.21)$$

Dabei werden folgende Annahmen getroffen:¹⁴⁵

1. Die stochastischen Nutzenkomponenten ε_w und $\varepsilon_{i|w}$ sind unabhängig voneinander.
2. Die stochastischen Nutzenkomponenten $\varepsilon_{i|w}$ sind unabhängig und identisch Gumbel-verteilt mit dem Skalierungsparameter μ_w .
3. ε_w ist so verteilt, dass der Zufallsnutzen U_{iw} der besten Alternative des Nests w (d. h. der Alternative mit dem größten Zufallsnutzen)

¹⁴² Vgl. *Louviere et al.* (2000), S. 144–145.

¹⁴³ Der Index zur Bezeichnung des Individuums n wird aus Gründen der besseren Lesbarkeit weggelassen.

¹⁴⁴ Vgl. *Louviere et al.* (2000), S. 145

¹⁴⁵ Vgl. *Ben-Akiva/Lerman* (1987), S. 286–287.

Gumbel-verteilt ist mit einem Skalierungsparameter, der mit λ_w bezeichnet wird.

Innerhalb eines Nests bleiben die Eigenschaften des multinomialen Logit-Modells erhalten, d. h. zwischen den Alternativen eines Nests besteht die IID- und die IIA-Eigenschaft, zwischen den Alternativen unterschiedlicher Nester jedoch nicht (notwendigerweise).¹⁴⁶ Dadurch werden zwischen den stochastischen Nutzenkomponenten der Alternativen eines Nests Korrelationen zugelassen, jedoch nicht zwischen den stochastischen Nutzenkomponenten der Alternativen unterschiedlicher Nester.¹⁴⁷ Unter diesen Annahmen ergibt sich die Wahlwahrscheinlichkeit der Alternative i im Nest w :¹⁴⁸

$$P_{wi} = \frac{e^{(\lambda_w V_w + \frac{\lambda_w}{\mu_w} IV_w)}}{\sum_{w' \in W} e^{(\lambda_{w'} V_{w'} + \frac{\lambda_{w'}}{\mu_{w'}} IV_{w'})}} \cdot \frac{e^{(\mu_w V_{i|w})}}{\sum_{j=1}^{J_w} e^{(\mu_w V_{j|w})}}, \quad (3.22)$$

mit

$$IV_w = \ln \sum_{j=1}^{J_w} e^{(\mu_w V_{j|w})}. \quad (3.23)$$

Der hintere Term in Gleichung (3.22) gibt die bedingte Wahrscheinlichkeit für die Wahl der Alternative i im Nest w an (dieser Term entspricht der Wahlwahrscheinlichkeit im multinomialen Logit-Modell); der vordere Term gibt die Wahrscheinlichkeit des bedingten Ereignisses, dass Alternativen des Nests w gewählt werden, an. Genau wie beim multinomialen Logit-Modell lässt sich die systematische Nutzenkomponente durch eine Nutzenfunktion ersetzen, beispielsweise der linear additiven Nutzenfunktion aus (3.3). Bei IV_w handelt es sich um den sogenannten Inklusivwert¹⁴⁹, welcher dazu verwendet wird, um die unterschiedlichen Ebenen miteinander zu verbinden; er entspricht dem Logarithmus des Nenners der bedingten Wahrscheinlichkeit. Bei dem dazugehörigen Inklusivwertpara-

¹⁴⁶ Vgl. Train (2009), S. 77–79.

¹⁴⁷ Vgl. Ben-Akiva/Lerman (1987), S. 286, Louviere et al. (2000), S. 144 und Train (2009), S. 79.

¹⁴⁸ Vgl. für die Herleitung der Wahlwahrscheinlichkeit Ben-Akiva/Lerman (1987), S. 285–293 und Train (2009), S. 79–86.

¹⁴⁹ Der Inklusivwert wird auch als Logsum bezeichnet, vgl. Hensher/Rose/Greene (2005), S. 487.

meter (*IV*-Parameter) handelt es sich um das Verhältnis der Skalierungsparameter zwischen der oberen und der unteren Ebene λ_w/μ_w .¹⁵⁰

Sofern der *IV*-Parameter im 0,1-Intervall liegt, gilt, dass das zugehörige Nested Logit-Modell konsistent mit der Nutzenmaximierung ist.¹⁵¹ Diese Eigenschaft ist dadurch bedingt, dass in den Varianzen der oberen Ebene auch die Varianzen der unteren Ebene enthalten sind. Da zwischen den Varianzen und den Skalierungsparametern ein umgekehrt proportionaler Zusammenhang besteht, gilt $\lambda_w \leq \mu_w$, bzw. $0 \leq \lambda_w/\mu_w \leq 1$.¹⁵²

Es ist nicht notwendig, dass ein Nest immer aus mehreren Alternativen besteht. Sofern sich unterhalb eines Zweigs nur eine einzige Alternative befindet, unterscheiden sich die Varianzen der beiden Ebenen nicht. Der Inklusivwertparameter beträgt dann 1.¹⁵³ In dem Fall, dass ein zweistufiges Nested Logit-Modell nur aus degenerierten Zweigen besteht oder aus anderen Gründen ausschließlich *IV*-Parameter von 1 ausweist, vereinfacht sich das NL-Modell zum multinomialen Logit-Modell.

In Analogie zum zweistufigen Nested Logit-Modell lassen sich Nested Logit-Modelle mit drei oder mehr Ebenen entwickeln.¹⁵⁴ Die 0,1-Restriktion der *IV*-Parameter gilt dabei auf allen Ebenen. Nicht nur degenerierte Zweige, sondern auch degenerierte Äste bzw. Stämme können existieren, unterhalb derer sich nur ein einziger Zweig bzw. Ast befindet. So, wie sich das zweistufige NL-Modell zu einem multinomialen Logit-Modell vereinfachen lässt, lassen sich auch NL-Modelle mit mehr Ebenen vereinfachen, unter der Voraussetzung, dass die *IV*-Parameter, die die jeweils obere Ebene mit der darunter liegenden Ebene verknüpfen, 1 betragen.

Im Gegensatz zum multinomialen Logit-Modell werden beim Nested Logit-Modell bzw. allgemein bei GEV-Modellen durch eine teilweise Lockerung der IID-Eigenschaft Korrelationen zwischen den nicht beobachteten Nutzenkomponenten mancher Alternativen möglich. Im

¹⁵⁰ Vgl. *Louviere et al.* (2000), S. 145–146 und *Hensher/Rose/Greene* (2005), S. 489.

¹⁵¹ Unter bestimmten Umständen kann das Nested Logit-Modell mit einem Inklusivwertparameter, der größer als 1 ist, konsistent mit der Nutzenmaximierung sein, vgl. dazu *Louviere et al.* (2000), S. 148 und die dort zitierten Quellen.

¹⁵² Vgl. *McFadden* (1979), S. 286–287, *McFadden* (1981), S. 229–230, *Ben-Akiva/Lerman* (1987), S. 288–289, *Louviere et al.* (2000), S. 146–147 und *Hensher/Rose/Greene* (2005), S. 490–494.

¹⁵³ Vgl. *Louviere et al.* (2000), S. 153–154 und *Hensher/Rose/Greene* (2005), S. 570–574.

¹⁵⁴ Vgl. *Ben-Akiva/Lerman* (1987), S. 291–295 und *Hensher/Rose/Greene* (2005), S. 496–515.

Gegensatz zu den Modellen, in denen die IID-Eigenschaft vollständig aufgegeben wird, ist keine Simulation¹⁵⁵ zur Bestimmung der Parameter erforderlich, sondern es kann stattdessen das Maximum Likelihood-Schätzverfahren eingesetzt werden.¹⁵⁶ Da die Log-Likelihood-Funktion eines Nested Logit-Modells nicht global konkav ist, handelt es sich bei einem gefundenen (relativen) Maximum nicht notwendigerweise auch um ein absolutes Maximum. Daher muss durch die Wahl von geeigneten Startwerten bzw. eines geeigneten Algorithmus sichergestellt werden, dass sich die Schätzwerte auf das absolute Maximum beziehen.¹⁵⁷

Da sich nicht alle Skalierungsparameter identifizieren lassen, ist es erforderlich, einen dieser Parameter auf 1 zu setzen, sodass die anderen Parameter relativ dazu bestimmt werden können. In der Regel wird entweder der Skalierungsparameter der untersten Ebene oder der obersten Ebene normalisiert. Die erstere Normalisierung wird auch als RU1, die letztere als RU2 bezeichnet.¹⁵⁸ Bei einer Normalisierung der obersten Skalierungsparameter ist sichergestellt, dass das Modell konsistent mit der Nutzenmaximierung ist, was bei einer Normalisierung der untersten Skalierungsparameter nicht ohne weitere Restriktionen der Fall ist.¹⁵⁹

Im Vorfeld ist es nicht immer eindeutig, welche Struktur bei einem Nested Logit-Modell zu verwenden ist, d. h. welche Alternativen in einem Nest zusammengefasst werden sollen. In diesem Fall ist es erforderlich, mehrere Modelle mit den unterschiedlichen Strukturen zu schätzen und das geeignetste Modell anhand eines Likelihood-Ratio-Tests (vgl. Kap. 3.5.2) auszuwählen.

¹⁵⁵ Vgl. *Train* (2001), S. 1–2 und *Train* (2009), S. 98.

¹⁵⁶ Vgl. *Louviere et al.* (2000), S. 144 und *Ben-Akiva/Lerman* (1987), S. 276, vgl. zum Maximum Likelihood-Schätzverfahren von Nested Logit-Modellen *Ben-Akiva/Lerman* (1987), S. 295–299.

¹⁵⁷ Vgl. *Train* (2009), S. 84–85.

¹⁵⁸ Die Bezeichnung „RU“ steht dabei für „Random Utility (Model)“, vgl. *Hensher/Rose/Greene* (2005), S. 496.

¹⁵⁹ Vgl. *Hunt* (2000), S. 99–101, *Hensher/Rose/Greene* (2005), S. 145–146, *Hensher/Greene* (2002) und *Hensher/Greene* (2002), S. 1–17.

3.4 Abgrenzung zu komplexeren Modellen der Discrete Choice Analyse

Zwar wird beim Nested Logit-Modell im Gegensatz zum multinomialen Logit-Modell die IIA-Eigenschaft teilweise gelockert, doch bleiben die Restriktionen bestehen, dass Nutzenparameter nicht zufällig zwischen den Individuen variieren können, um Heterogenität abzubilden. Außerdem ist die für die Analyse von Paneldaten erforderliche Abbildung von Korrelationen der stochastischen Nutzenkomponenten nicht möglich.¹⁶⁰ Beim multinomialen Probit-Modell bestehen diese Restriktionen nicht. Hier wird für alle stochastischen Nutzenkomponenten eine Normalverteilung unterstellt. Das Integral der zugehörigen Wahrscheinlichkeitsfunktion dieses Modells kann nicht direkt bestimmt werden und erfordert daher eine Simulation.

Die Restriktion der Verhaltenshomogenität besteht beim sogenannten Mixed Logit-Modell ebenfalls nicht mehr. Hier wird der Nutzenparameter in eine systematische und eine stochastische Komponente zerlegt.¹⁶¹ Für die stochastische Komponente können unterschiedliche Verteilungen angenommen werden können. Die Bestimmung der Parameter erfordert eine Simulation, beispielsweise mittels Maximum Simulated Likelihood oder Hierarchical Bayes.¹⁶² Bei der Verwendung von letzterem Verfahren wird das Modell auch als Hierarchical Bayes Modell bezeichnet.¹⁶³ Wenn als Verteilungsfunktion beim Mixed Logit-Modell eine diskrete Verteilung verwendet wird, geht das Mixed Logit Modell ins sogenannte Latent Class Modell über.¹⁶⁴ Hierbei wird angenommen, dass es eine geringe Anzahl von latenten, d. h. nicht beobachtbaren Klassen mit unterschiedlichen Nutzenparametern gibt. Ein Individuum wird den unterschiedlichen Klassen mit einer bestimmten Wahrscheinlichkeit zugeordnet. Sowohl beim Mixed Logit- als auch beim Latent Class Modell besteht die Möglichkeit, individuen-spezifische Parameter zu bestimmen, die jedoch auf aggregierten Daten basieren.¹⁶⁵ Bei einer weiteren Verallgemeinerung

¹⁶⁰ Vgl. *Train* (2009), S. 97.

¹⁶¹ Vgl. *Hensher/Rose/Greene* (2005), S. 606–610.

¹⁶² Vgl. *Train* (2001), S. 2–5.

¹⁶³ Vgl. *Allenby/Ginter* (1995), S. 393.

¹⁶⁴ Vgl. *Train* (2009), S. 135.

¹⁶⁵ Vgl. *Train* (2009), S. 259–264.

der Modelle lässt sich nicht nur die Heterogenität der Nutzenparameter, sondern auch die der Skalierungsparameter abbilden.¹⁶⁶

3.5 Testverfahren

3.5.1 Hausman-McFadden-Test der Unabhängigkeit von irrelevanten Alternativen

Die Unabhängigkeit von irrelevanten Alternativen wird bei der Substitution von Alternativen deutlich. Im Extremfall, wie beim „Red Bus/Blue Bus“-Paradoxon gezeigt (vgl. Kap. 3.2), lässt sich eine Alternative vollständig durch eine andere Alternative substituieren. Da jedoch in Discrete Choice Experimenten in der Regel nur die gewählte Alternative erfragt wird, ist nicht direkt bekannt, durch welche Alternative/n eine Alternative substituiert würde.

Ein Testverfahren zur Überprüfung der IIA-Annahme wurde von *Hausman/McFadden* (1984) vorgeschlagen.¹⁶⁷ Dabei wird ein unrestringiertes multinomiales Logit-Modell, bei dem das vollständige Choice Set verwendet wird, mit einem restringierten multinomialen Logit-Modell, bei dem eine oder mehrere Alternative/n aus dem Choice Set ausgeschlossen werden, verglichen. Zur Bestimmung des restringierten Modells müssen entsprechend die nicht verfügbare/n Alternative/n sowie alle Wahlsituationen, in denen diese Alternative/n gewählt wurden, ausgeschlossen werden.¹⁶⁸ Das restringierte Modell weist daher weniger Wahlsituationen auf als das unrestringierte Modell. Da die Wahlwahrscheinlichkeiten in den Modellen von den Nutzenparametern abhängen, ist es beim Vergleich der beiden Modelle ausschlaggebend, dass sich die Nutzenparameter bei einer Veränderung des Choice Sets nicht ändern. Bei dem Vergleich werden die in beiden Modellen verwendeten Nutzenparameter in der Regel ohne die alternativenspezifischen Konstanten¹⁶⁹ berücksichtigt, d. h.

¹⁶⁶ Vgl. *Fiebig et al.* (2010) für das sogenannte „Generalized Multinomial Logit Modell“ sowie *Magidson/Vermunt* (2008) für das sogenannte „Scale Adjusted Latent Class Modell“

¹⁶⁷ Vgl. zu anderen Testverfahren *Ben-Akiva/Lerman* (1987), S. 184–194 und die dort zitierte Literatur.

¹⁶⁸ Vgl. *Ben-Akiva/Lerman* (1987), S. 184 und *Hensher/Rose/Greene* (2005), S. 597–599.

¹⁶⁹ Vgl. *Hensher/Rose/Greene* (2005), S. 597.

spezifische Nutzenparameter der ausgeschlossenen Alternative/n werden nicht berücksichtigt.

Die Teststatistik des Hausman-McFadden-Test der IIA lautet:¹⁷⁰

$$T = (\boldsymbol{\beta}_r - \boldsymbol{\beta}_u)' [\text{cov}(\boldsymbol{\beta}_r) - \text{cov}(\boldsymbol{\beta}_u)]^{-1} (\boldsymbol{\beta}_r - \boldsymbol{\beta}_u), \quad (3.24)$$

wobei es sich bei $\boldsymbol{\beta}_r$ und $\boldsymbol{\beta}_u$ um Vektoren der Nutzenparameter und bei $\text{cov}(\boldsymbol{\beta}_r)$ und $\text{cov}(\boldsymbol{\beta}_u)$ um Kovarianzmatrizen des restringierten (r) bzw. des unrestringierten (u) Modells handelt. Das Testergebnis wird mit dem Wert der Chi-Quadrat-Verteilung verglichen, wobei die Anzahl der Freiheitsgrade der Anzahl der im restringierten Modell verwendeten Nutzenparameter entspricht. Sofern der kritische Wert nicht überschritten wird, bedeutet dies, dass die Annahme der Unabhängigkeit von irrelevanten Alternativen für das gegebene unrestringierte Modell nicht abgelehnt wird. Dahingegen bedeutet eine Überschreitung des kritischen Werts, dass entweder die IIA-Annahme abgelehnt wird, eine Fehlspezifikation der Attribute des Modells vorliegt oder beides gleichzeitig.¹⁷¹

3.5.2 Likelihood-Ratio-Test

Mithilfe des Likelihood-Ratio-Tests kann insbesondere überprüft werden, ob sich ein Modell signifikant verbessert, wenn ein oder mehrere weitere/r Parameter in die Nutzenfunktion aufgenommen werden oder, ob sich die Nutzenfunktionen zwischen Gruppen von Individuen unterscheiden. Dabei handelt es sich um einen aggregierten Test, bei dem der Log-Likelihood-Wert eines restringierten Modells mit dem Log-Likelihood-Wert eines unrestringierten Modells bzw. der Summe der Log-Likelihood-Werte mehrerer unrestringierter Modelle verglichen wird. Wird z. B. getestet, ob ein weiterer Parameter in die Nutzenfunktion aufgenommen wird, besteht beim restringierten Modell die Restriktion dahingehend, dass der entsprechende Parameter 0 gesetzt wird, was bedeutet, dass dieser Parameter nicht geschätzt wird; im unrestringierten Modell wird hingegen für diesen Parameter ein Wert geschätzt. Wird getestet, ob Unterschiede zwischen den Nutzenfunktionen von Gruppen von Indivi-

¹⁷⁰ Vgl. zur Herleitung der Teststatistik *Hausman/McFadden* (1984), S. 1221–1226.

¹⁷¹ Vgl. *Hausman/McFadden* (1984), S. 1224.

duen vorliegen, besteht im restringierten Modell die Restriktion der Gleichheit der Nutzenfunktionen dadurch, dass für die Gruppen ein gemeinsames Modell mit gemeinsamen Parametern geschätzt wird, während die Gleichheitsrestriktionen durch die Schätzung von getrennten Modellen für die jeweiligen Gruppen nicht besteht. Unter der Annahme, dass die Restriktion wahr ist, beträgt die Teststatistik¹⁷²

$$\lambda = -2[LL_R - \sum_{k=1}^K LL_{U_k}], \quad (3.25)$$

wobei es sich bei LL_R um den Log-Likelihood-Wert des restringierten Modells und bei LL_{U_k} um den Log-Likelihood-Wert des k -ten unrestringierten Modells handelt bei insgesamt K unrestringierten Modellen. Das Testergebnis ist asymptotisch χ^2 -verteilt¹⁷³ mit der Anzahl von Freiheitsgraden, welcher der Differenz der Anzahl der Parameter in dem/den unrestringierten Modell/en und dem restringierten Modell entspricht. Überschreitet das Testergebnis den kritischen Wert nicht, wird das restringierte Modell beibehalten, andernfalls wird das restringierte Modell zu Gunsten des unrestringierten Modells/der unrestringierten Modelle verworfen.

Neben der Überprüfung, ob ein einzelner Parameter in ein Modell aufgenommen werden soll, insbesondere dergestalt, dass ein alternativenspezifischer Parameter anstelle eines generischen Parameters (d. h. ein für alle Alternativen identischer Parameter) verwendet wird, wird der Likelihood-Ratio-Test auch zur Überprüfung der Gesamtsignifikanz eines Modells verwendet. Dabei wird das zu überprüfende Modell (unrestringiertes Modell, d. h. die Nutzenparameter sind frei schätzbar) mit dem sogenannten Konstantenmodell¹⁷⁴ verglichen.¹⁷⁵ Gleichmaßen ist ein

¹⁷² Dabei wird auch angenommen, bei dem restringierten Modell handelt es sich um das „wahre“ Modell. Wird beispielsweise ein multinomiales Logit-Modell verwendet und die IIA-Annahme ist verletzt, so ist das Testverfahren ungeeignet, da das MNL-Modell nicht das „wahre“ Modell ist, vgl. *Swait/Louviere* (1993), S. 307.

¹⁷³ Vgl. für den zugehörigen Beweis *Theil* (1971), S. 396–397.

¹⁷⁴ In einem Konstantenmodell werden nur Nutzenparameter für die alternativenspezifischen Konstanten geschätzt. Entsprechend sind alle anderen Parameter auf 0 restringiert, d. h. nicht enthalten.

¹⁷⁵ Vgl. *Louviere et al.* (2000), S. 53–54.

Vergleich eines geschätzten Modells mit dem sogenannten Nullmodell¹⁷⁶ möglich.

Im multinomialen Logit-Modell lässt sich der Skalierungsparameter nicht identifizieren. Ebenso lassen sich im Nested Logit-Modell nicht alle Skalierungsparameter identifizieren. Daher handelt es sich bei einem geschätzten Parameter im eigentlichen Sinne nicht um einen Nutzenparameter, sondern um das Produkt aus Skalierungsparameter und Nutzenparameter; die Nutzen- und Skalierungsparameter sind miteinander „konfundiert“, d. h. sie sind miteinander vermischt. Wird der oben beschriebene (einstufige) Likelihood-Ratio-Test in (3.25) zum Vergleich der Nutzenfunktionen unterschiedlicher Gruppen verwendet, so kann, wenn die Annahme der Gleichheit der Nutzenfunktionen abgelehnt wird, nicht unterschieden werden, ob der Unterschied aus den Skalierungsparametern und/oder den Nutzenparametern resultiert.¹⁷⁷

Mit dem zweistufigen Likelihood-Ratio-Test¹⁷⁸ ist es möglich, zu überprüfen, ob sich sowohl die Nutzenparameter als auch die Skalierungsparameter zwischen zwei Modellen¹⁷⁹ nicht unterscheiden (d. h. $\beta_1 = \beta_2$ und $\mu_1 = \mu_2$), ob die Nutzenparameter zwar übereinstimmen, jedoch Unterschiede zwischen den Skalierungsparametern bestehen (d. h. $\beta_1 = \beta_2$ und $\mu_1 \neq \mu_2$) oder ob zwischen den Modellen „echte“ Unterschiede bestehen (d. h. $\mu_1\beta_1 \neq \mu_2\beta_2$).¹⁸⁰ Da die Skalierungsparameter und die Nutzenparameter konfundiert sind, lässt sich bei „echten“ Unterschieden keine Aussage darüber treffen, ob sich sowohl Nutzenparameter als auch

¹⁷⁶ In einem Nullmodell sind keine Nutzenparameter enthalten. Entsprechend sind alle Parameter, auch die der alternativenspezifischen Konstanten, auf 0 restringiert, d. h. nicht enthalten. In diesem Modell ist die Wahlwahrscheinlichkeit für alle Alternativen gleich groß.

¹⁷⁷ Relative Nutzenparameter wie beispielsweise Zahlungsbereitschaften sind gleich, wenn sich die Nutzenparameter zwischen den Gruppen nicht unterscheiden, unabhängig davon, ob hinsichtlich der Skalierungsparameter Unterschiede bestehen oder nicht. Insofern ist es von Interesse, eine Differenzierung in Bezug auf Skalierungs- und Nutzenparameter vorzunehmen.

¹⁷⁸ Vgl. *Swait/Louviere* (1993), S. 306–309.

¹⁷⁹ Aus Übersichtlichkeitsgründen beschränken sich die folgenden Ausführungen zunächst auf den Vergleich von zwei von der Struktur her identischen multinomialen Logit-Modellen. Der Vergleich von unterschiedlichen Modellen (beispielsweise ein Vergleich von einem Modell mit zwei Alternativen mit einem anderen Modell, das sich hinsichtlich einer dritten Alternative unterscheidet), von mehr als zwei Modellen, sowie von Nested Logit-Modellen ist jedoch möglich.

¹⁸⁰ Vgl. *Swait/Louviere* (1993), S. 307.

Skalierungsparameter unterscheiden (d. h. $\beta_1 \neq \beta_2$ und $\mu_1 \neq \mu_2$) oder ob sich lediglich die Nutzenparameter unterscheiden, die Skalierungsparameter hingegen gleich sind (d. h. $\beta_1 \neq \beta_2$ und $\mu_1 = \mu_2$). Bei dem zweistufigen Testverfahren soll die Hypothese überprüft werden, dass

$$H_1: \beta_1 = \beta_2 \text{ und } \mu_1 = \mu_2 \quad (3.26)$$

gilt. Diese Hypothese wird anhand von zwei Subhypothesen, bezeichnet mit H_{1A} und H_{1B} , überprüft. In der ersten Stufe des zweistufigen Likelihood-Ratio-Tests wird die Subhypothese überprüft, ob sich die Nutzenparameter zwischen den Modellen nicht unterscheiden, wobei mögliche Unterschiede zwischen den Skalierungsparametern zugelassen werden. Die erste Subhypothese lautet entsprechend

$$H_{1A}: \beta_1 = \beta_2. \quad (3.27)$$

Sofern diese Subhypothese abgelehnt wird, wird automatisch auch die Gesamthypothese H_1 abgelehnt. Andernfalls folgt die Überprüfung der zweiten Subhypothese, deren Gegenstand die Gleichheit der Skalierungsparameter ist:

$$H_{1B}: \mu_1 = \mu_2. \quad (3.28)$$

Zur Überprüfung der ersten Subhypothese ist die Schätzung eines künstlichen Nested Logit-Modells¹⁸¹ erforderlich, welches dazu verwendet wird, um mögliche Unterschiede zwischen den Skalierungsparametern der beiden Modelle zuzulassen. Auf der obersten Ebene, d. h. der Zweigebene, wird zwischen den beiden Modellen unterschieden. Darunter befinden sich die Alternativen des ersten Modells einerseits sowie die Alternativen des zweiten Modells andererseits. Die Anzahl der elementaren Alternativen hat sich dadurch verdoppelt. Das Nested Logit-Modell wird als „künstlich“ bezeichnet, weil die Alternativen nicht als wählbare Alternativen verfügbar sind, sondern ausschließlich zur Berücksichtigung möglicher Varianzunterschiede verwendet werden.

¹⁸¹ Aus diesem Grund wird der zweistufige Likelihood-Ratio-Test auch als „Nested Logit Trick“ bezeichnet.

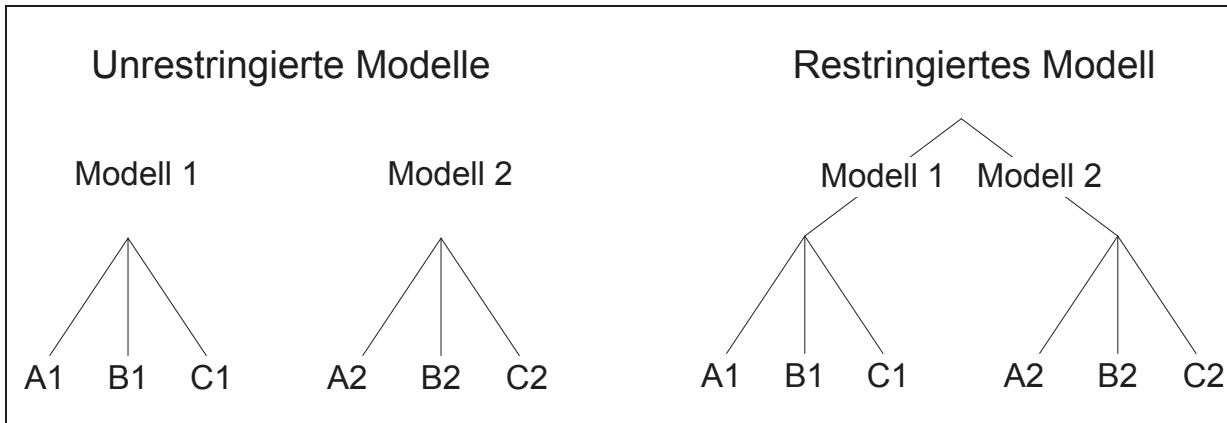
Bei diesem künstlichen Nested Logit-Modell handelt es sich um das restringierte Modell, wobei die Restriktion darin besteht, dass für die Alternativen der beiden unterschiedlichen Modelle gemeinsame (d. h. auf Gleichheit restringierte) Nutzenparameter geschätzt werden. Zudem wird einer der Inklusivwertparameter auf 1 gesetzt, die Schätzung des anderen *IV*-Parameters erfolgt zur Abbildung möglicher Unterschiede zwischen den Skalierungsparametern relativ. Darüber hinaus wird für jede der beiden Gruppen ein einzelnes multinomiales Logit-Modell geschätzt. Hinsichtlich der Nutzenparameter, oder genauer gesagt, der mit dem jeweiligen Skalierungsparameter konfundierten Nutzenparameter, liegen keine Restriktionen vor. Exemplarisch ist die Struktur der unrestringierten Modelle und des restringierten Modells in Abb. 3.2 dargestellt. Die Teststatistik zur Überprüfung der ersten Subhypothese lautet:

$$\lambda_A = -2[LL_{ges\ \mu} - \sum_{k=1}^K LL_k], \quad (3.29)$$

wobei es sich bei $LL_{ges\ \mu}$ um den Log-Likelihood-Wert des gemeinsamen Modells handelt, bei dem die Nutzenparameter auf Gleichheit restringiert wurden, aber mögliche Unterschiede zwischen den Skalierungsparametern zugelassen werden, und bei LL_k um den Log-Likelihood-Wert des k -ten unrestringierten Modells. Bei dem Vergleich von zwei Modellen gilt $K = 2$. Das Testergebnis ist asymptotisch χ^2 -verteilt mit der Anzahl der Freiheitsgrade, welche der Differenz der Anzahl der (mit den Skalierungsparametern konfundierten) Nutzenparameter der einzelnen Modelle und der Anzahl der Nutzenparameter sowie des relativen Skalierungsparameters (in Form des *IV*-Parameters) des gemeinsamen, restringierten Modells entspricht. Wird der kritische Wert überschritten, erfolgt eine Ablehnung der ersten Subhypothese sowie automatisch eine Ablehnung der Gesamthypothese. Wird hingegen der kritische Wert nicht überschritten, kann die Annahme der Gleichheit der Nutzenparameter nicht abgelehnt werden und es erfolgt eine Überprüfung der zweiten Subhypothese, der Gleichheit der Skalierungsparameter.

Zu diesem Zweck ist die Schätzung eines gemeinsamen Modells erforderlich, bei dem die Skalierungsparameter der unterschiedlichen Gruppen auf Gleichheit restringiert werden. Die Restriktion der Gleichheit

Abb. 3.2: Exemplarische Struktur der unrestringierten Modelle und des restringierten Modells zur Überprüfung der ersten Subhypothese des zweistufigen Likelihood-Ratio-Tests



Quelle: Eigene Darstellung.

der Skalierungsparameter ist modellimmanent beim multinomialen Logit-Modell gegeben.¹⁸² Dieses Modell entspricht einem zweistufigen NL-Modell, bei dem die *IV*-Parameter auf Gleichheit restringiert werden. Das restringierte Modell wird mit dem unrestringierten Modell, bei dem mögliche Unterschiede zwischen den Skalierungsparametern der unterschiedlichen Gruppen zugelassen werden (bei diesem Modell handelt es sich um das restringierte Modell der Überprüfung der ersten Subhypothese), verglichen. Exemplarisch ist die Struktur des unrestringierten und des restringierten Modells in Abb. 3.3 dargestellt.

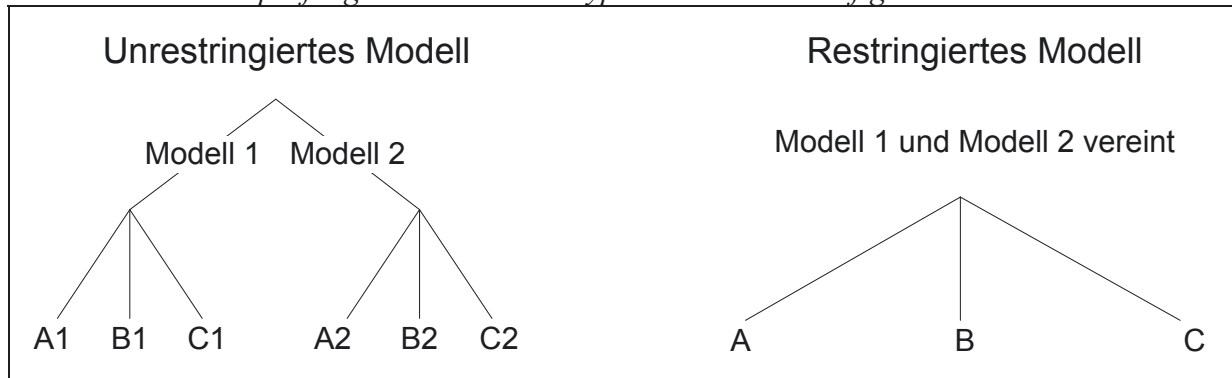
Die Teststatistik zur Überprüfung der zweiten Subhypothese lautet:

$$\lambda_B = -2[LL_{ges} - LL_{ges \mu}], \quad (3.30)$$

wobei es sich bei LL_{ges} um den Log-Likelihood-Wert des restringierten Modells und bei $LL_{ges \mu}$ um den Log-Likelihood-Wert des unrestringierten Modells handelt. Das Testergebnis ist asymptotisch χ^2 -verteilt mit der Anzahl der Freiheitsgrade, welche der Differenz der Anzahl der Nutzenparameter und des relativen Skalierungsparameters (in Form des *IV*-Parameters) im unrestringierten Modell und der Anzahl der Nutzenparameter im restringierten Modell entspricht. Sofern der kritische Wert nicht überschritten wird, kann die zweite Subhypothese und somit auch die Gesamthypothese nicht abgelehnt werden. Es liegt somit sowohl

¹⁸² Vgl. Hensher/Rose/Greene (2005), S. 489.

Abb. 3.3: Exemplarische Struktur des unrestringierten und des restringierten Modells zur Überprüfung der zweiten Subhypothese des zweistufigen Likelihood-Ratio-Tests



Quelle: Eigene Darstellung.

Gleichheit der Nutzenparameter als auch der Skalierungsparameter vor. Wird der kritische Wert hingegen überschritten, bedeutet dies, dass die Nutzenparameter gleich sind, jedoch Unterschiede zwischen den Skalierungsparametern bestehen.

Dieses zweistufige Testverfahren ist nicht auf den Vergleich von nur zwei Gruppen bzw. Modellen¹⁸³ beschränkt. Wenn die Gleichheit von mehr als zwei Gruppen bzw. Modellen miteinander überprüft werden soll, ist es erforderlich, bei dem restringierten Modell zur Überprüfung der ersten Subhypothese weitere Zweige mit relativen *IV*-Parametern für die zusätzlichen Gruppen hinzuzufügen. Sofern jedoch bei einem gemeinsamen Vergleich von mehr als zwei Gruppen Unterschiede zwischen den Nutzenparametern und/oder Skalierungsparametern festgestellt werden, ist anhand des beschriebenen Testverfahrens nicht ersichtlich, ob Unterschiede bei allen oder nur bei manchen Gruppen bestehen, sodass in diesem Fall paarweise Vergleiche geeigneter erscheinen.

Zudem ist es möglich, das zweistufige Testverfahren auch dann einzusetzen, wenn sich die Modelle teilweise hinsichtlich des Choice Sets und/oder der verwendeten Attribute (und damit hinsichtlich der Nutzenparameter) unterscheiden. Der Vergleich bezieht sich dann auf die gemeinsamen Alternativen und gemeinsamen Attribute.

Handelt es sich bei den zu vergleichenden Modellen um zweistufige NL-Modelle, ist für die Überprüfung der ersten Subhypothese die Schätzung

¹⁸³ Dabei handelt es sich bei dem Testverfahren immer um einen Vergleich zwischen restringierten und unrestringierten Modellen, vgl. *Ben-Akiva/Lerman* (1987), S. 171.

eines dreistufigen NL-Modells erforderlich.¹⁸⁴ Dabei wird auf der Ebene der Äste zwischen den Modellen unterschieden, unterhalb dessen sich die jeweiligen zu testenden Modelle befinden. Bei diesem gemeinsamen Modell kann unterschieden werden, ob die *IV*-Parameter auf der Ebene der Zweige auf Gleichheit restringiert werden oder nicht. Bei dem restringierten Modell zur Überprüfung der zweiten Subhypothese handelt es sich um ein zweistufiges NL-Modell, welches einem dreistufigen NL-Modell, bei dem die *IV*-Parameter der Astebene auf Gleichheit restringiert sind, entspricht.

Der zweistufige Likelihood-Ratio-Test wird insbesondere zur Überprüfung eingesetzt, ob Modelle, die auf Realdaten basieren, mit Modellen, die auf Experimentaldaten basieren, zu Gesamtmodellen vereint werden können.¹⁸⁵ Zudem wird dieser Test zum Vergleich zwischen Modellen, deren Daten zu unterschiedlichen Zeitpunkten und/oder an unterschiedlichen Orten erhoben wurden¹⁸⁶, sowie zur Analyse von Reihenfolge- oder Ermüdungseffekten¹⁸⁷ eingesetzt.

3.5.3 Modellgüte und Validität

Üblicherweise wird das Maximum Likelihood-Schätzverfahren (vgl. Kap. 3.2) zur Bestimmung der Nutzenparameter eines MNL- oder eines NL-Modells eingesetzt. Bei der Modellschätzung ist a priori festzulegen, welche Attribute in die Nutzenfunktionen aufgenommen werden sollen und infolgedessen, für welche (Nutzen-)Parameter ein Wert geschätzt werden soll. Die Güte des geschätzten Modells kann sowohl in Bezug auf die Gesamtsignifikanz als auch in Bezug auf die Signifikanz einzelner Parameter beschrieben werden.

Zur Überprüfung der Gesamtsignifikanz eines Modells kann der Likelihood-Ratio-Test eingesetzt werden.¹⁸⁸ Dabei ist neben der Schätzung des zu überprüfenden Modells auch die Schätzung eines Konstantenmodells erforderlich, in dem nur die alternativenspezifischen Konstanten enthalten

¹⁸⁴ Vgl. *Swait/Bernardino* (2000), S. 4–9.

¹⁸⁵ Vgl. *Hensher/Bradley* (1993), *Bradley/Daly* (1997) und *Hensher/Louviere/Swait* (1998).

¹⁸⁶ Vgl. *Swait/Bernardino* (2000) und *Severin/Louviere/Finn* (2001).

¹⁸⁷ Vgl. *Bradley/Daly* (1994) und *Bradley/Daly* (1992).

¹⁸⁸ Vgl. *McFadden* (1974), S. 120–121, *Louviere et al.* (2000), S. 53–54, *Train* (2009), S. 70, *Hensher/Rose/Greene* (2005), S. 330–332 und *Ben-Akiva/Lerman* (1987), S. 164–166.

sind, was einer Restriktion der Nutzenparameter auf 0 entspricht. Sofern das zu überprüfende Modell einen signifikant besseren Log-Likelihood-Wert als das Konstantenmodell aufweist, ist die Gesamtsignifikanz des Modells gegeben.

Eng verbunden mit dem Likelihood-Ratio-Test zum Vergleich eines Modells mit dem Konstantenmodell ist der sogenannte Likelihood-Ratio-Index, welcher für ein zu testendes Modell definiert ist als

$$\rho^2 = 1 - \frac{LL(\beta)}{LL(const.)}, \quad (3.31)$$

wobei es sich bei $LL(\beta)$ um den Log-Likelihood-Wert des zu testenden Modells und bei $LL(const.)$ um den Log-Likelihood-Wert des zugehörigen Konstantenmodells handelt.¹⁸⁹ Der Log-Likelihood-Wert des zu überprüfenden Modells kann nicht schlechter, d. h. geringer sein als der des Konstantenmodells.¹⁹⁰ Im ungünstigsten Fall weisen die Nutzenparameter keinen Erklärungsgehalt auf, sodass sie 0 betragen. Zwischen den beiden Log-Likelihood-Werten würde dann kein Unterschied bestehen, sodass $LL(\beta) = LL(const.)$ bzw. $LL(\beta)/LL(const.) = 1$ gilt. Daraus ergibt sich die untere Begrenzung des Likelihood-Ratio-Indexes mit $\rho^2 = 0$. Im günstigsten Fall werden die Wahlentscheidungen durch das Modell perfekt abgebildet, sodass die Wahlwahrscheinlichkeit für eine gewählte Alternative jeweils exakt 1 beträgt und daher zwischen den Wahlwahrscheinlichkeiten und den Wahlentscheidungen kein Unterschied besteht. In diesem Fall würde der Log-Likelihood-Wert 0 betragen, woraus sich die obere Begrenzung des Likelihood-Ratio-Indexes mit $\rho^2 = 1$ ergibt.

Zwar wird der Likelihood-Ratio-Index ρ^2 auch als Pseudo- R^2 oder McFadden's (Pseudo) R^2 bezeichnet, jedoch beschränkt sich die Gemeinsamkeit mit dem Bestimmtheitsmaß R^2 einer Regressionsanalyse auf den 0,1-Wertebereich.¹⁹¹ Während das R^2 bei einer linearen Regression den Anteil der erklärten Varianz an der Gesamtvarianz angibt, existiert für den

¹⁸⁹ Vgl. *McFadden* (1974), S. 120–121, *Train* (2009), S. 68–69, *Ben-Akiva/Lerman* (1987), S. 167 und *Louviere et al.* (2000), S. 54.

¹⁹⁰ Dabei besteht die Voraussetzung, dass in dem zu überprüfenden Modell zumindest die alternativenspezifischen Konstanten enthalten sind.

¹⁹¹ Vgl. *Train* (2009), S. 68–69, *Hensher/Louviere/Swait* (1998), S. 337 und *Louviere et al.* (2000), S. 54.

Likelihood-Ratio-Index keine analoge Interpretationsmöglichkeit. Der Likelihood-Ratio-Index kann für den Vergleich von unterschiedlichen Modellen nur dann herangezogen werden, wenn zur Modellierung die gleichen Wahldaten und die gleichen Alternativen verwendet werden. Dabei bezieht sich der Vergleich darauf, dass das Modell mit dem größeren Likelihood-Ratio-Index als das bessere Modell gilt. Allerdings würde ein Vergleich der Log-Likelihood-Werte zum gleichen Ergebnis führen. Sobald bei den Modellen unterschiedliche Daten und/oder Alternativen verwendet werden, ist kein Vergleich mittels Likelihood-Ratio-Index möglich. Diese Problematik besteht auch bei dem adjustierten Likelihood-Ratio-Index¹⁹², welcher um die Anzahl der verwendeten Parameter und ggf. die Anzahl der Wahlalternativen korrigiert wird, fort.

In manchen Fällen wird beim Likelihood-Ratio-Test und beim Likelihood-Ratio-Index anstelle des Konstantenmodells das Nullmodell verwendet. Da beim Nullmodell alle Alternativen die gleiche Wahlwahrscheinlichkeit aufweisen, oftmals aber keine solche Gleichverteilung bei der Wahl der Alternativen vorliegt, führt in der Regel bereits die Berücksichtigung von alternativenspezifischen Konstanten zu einer deutlichen Verbesserung des Log-Likelihood-Werts; eine Aussage hinsichtlich der Gesamtsignifikanz des Modells ist daher entsprechend gering und ein Vergleich mit dem Konstantenmodell zu bevorzugen.¹⁹³

Mithilfe eines *t*-Tests lässt sich die Signifikanz eines einzelnen Parameters überprüfen, indem getestet wird, ob der einzelne Parameter von einem bestimmten Wert abweicht.¹⁹⁴ Dabei ist insbesondere von Bedeutung, ob sich ein Nutzenparameter signifikant von 0 und ob sich ein *IV*-Parameter signifikant von 1 unterscheidet. Darüber hinaus lässt sich überprüfen, ob die signifikanten Parameter plausible Vorzeichen aufweisen (Augenscheinvalidität) sowie ob die relativen Parameter, beispielsweise in Form von Zahlungsbereitschaften, plausibel erscheinen. Bei letzterem ist auch ein Vergleich mit den Ergebnissen anderer Studien möglich.¹⁹⁵

¹⁹² Vgl. *Ben-Akiva/Lerman* (1987), S. 167 und *Louviere et al.* (2000), S. 55.

¹⁹³ Vgl. *Louviere et al.* (2000), S. 54–55.

¹⁹⁴ Vgl. *Train* (2009), S. 70 und *Ben-Akiva/Lerman* (1987), S. 161.

¹⁹⁵ Vgl. *Ben-Akiva/Lerman* (1987), S. 157–160.

Als Indikator, ob das Modell die verwendeten Daten geeignet abbildet (interne Validität), wird ein Verfahren vorgeschlagen, bei dem der Anteil der korrekt vorhergesagten Wahlentscheidungen bestimmt wird.¹⁹⁶ Die dabei verwendete Heuristik besteht darin, dass für jede einzelne Wahlentscheidung die anhand des Modells berechnete Wahlwahrscheinlichkeit mit der jeweiligen Wahlentscheidung im Experiment bzw. der realen Entscheidungssituation verglichen wird. Weist die tatsächlich gewählte Alternative die höchste Wahlwahrscheinlichkeit auf, wird die Wahl als korrekt vorhergesagt angesehen. Auf diese Weise lässt sich insgesamt der Anteil der korrekt vorhergesagten Wahlentscheidungen¹⁹⁷ bestimmen. Allerdings ist dieses Vorgehen nicht frei von Kritik, welche hauptsächlich darin besteht, dass das Verfahren dem Sinn der Wahlwahrscheinlichkeiten nicht entspricht. Bedingt dadurch, dass nicht der Anspruch erhoben wird, dass alle Einflüsse im Modell abgebildet werden (vgl. Kap. 3.1), erfolgt die Verwendung des Zufallsnutzenkonzepts, welches dazu führt, dass eine Wahlentscheidung nicht exakt, sondern nur mit einer bestimmten Wahrscheinlichkeit vorhergesagt werden kann. Wenn jedoch aus der größten Wahlwahrscheinlichkeit eine einzelne Wahlentscheidung vorhergesagt wird, ist dies gleichbedeutend mit der Annahme, im Modell würden alle bestehenden Einflüsse vollständig und somit rein deterministisch abgebildet.

Zudem ist eine Vorhersage mit diesem heuristischen Verfahren mit einer Verzerrung verbunden, welche am folgenden Beispiel¹⁹⁸ veranschaulicht werden soll: Angenommen, die aggregierten Wahlwahrscheinlichkeiten der Alternativen A und B werden in einem Modell perfekt wiedergeben und betragen $P(A) = 0,75$ und $P(B) = 0,25$. Eine optimale Prognose käme zu dem Ergebnis, dass von 100 Befragten insgesamt 75 die Alternative A und 25 die Alternative B wählen würden.¹⁹⁹ Eine derartige optimale Prognose ist jedoch nicht individuen-spezifisch möglich. Eine Prognose unter Verwendung des beschriebenen heuristischen Verfahrens ergäbe, dass die Alternative A in allen 100 Wahlentscheidungen gewählt würde, die Alternative B hingegen nie. Diese Verzerrung wird noch deutlicher,

¹⁹⁶ Vgl. *Hensher/Rose/Greene* (2005), S. 382–383 und *Ben-Akiva/Lerman* (1987), S. 91–92.

¹⁹⁷ Der Anteil der korrekt vorhergesagten Wahlentscheidungen an allen Wahlentscheidungen wird auch als Trefferquote oder „Hit-Rate“ bezeichnet.

¹⁹⁸ Vgl. *Train* (2009), S. 69.

¹⁹⁹ Vgl. *Train* (2009), S. 31 und *Hensher/Rose/Greene* (2005), S. 366.

wenn sich die Wahlwahrscheinlichkeiten der Alternativen im Modell nur unwesentlich unterscheiden, beispielsweise wenn die Wahlwahrscheinlichkeiten $P(A) = 0,51$ und $P(B) = 0,49$ betragen.

3.6 Zusammenfassung

Bei der Discrete Choice Analyse handelt es sich um ein dekompositionelles Verfahren zur Analyse von Nutzenstrukturen. Dabei wählt ein Individuum aus mehreren Wahlalternativen, die durch Kombinationen von Attributsausprägungen charakterisiert werden, genau eine Alternative aus. Derartige Wahlentscheidungen können als Real- oder Experimentaldaten erhoben werden. Durch eine Einbeziehung von Realdaten in Wahlexperimente können ebendiese realistischer gestaltet und ein Bezug zu einer tatsächlichen Handlung hergestellt werden.

Im Rahmen der Analyse wird angenommen, dass ein Individuum genau die Alternative wählt, deren Attributsausprägungen ihm den größten Nutzen stiften. Unter Verwendung des Zufallsnutzenkonzepts wird der Gesamtnutzen einer Alternative in eine deterministische und eine stochastische Nutzenkomponente zerlegt. Mit der stochastischen Nutzenkomponente wird nicht ausgedrückt, dass sich das Individuum bei seinen Entscheidungen unsicher ist oder seine Entscheidungen gar zufällig trifft, sondern, dass nicht der Anspruch erhoben wird, dass alle bestehenden Einflüsse im Modell explizit abgebildet werden können.

Unter der Annahme, dass die stochastischen Nutzenkomponenten voneinander unabhängig, identisch und Gumbel-verteilt sind, ergibt sich das multinomiale Logit (MNL)-Modell und dessen Sonderfall für zwei Wahlalternativen, das binomiale Logit-Modell. Mit dieser Verteilungsannahme ist jedoch die Annahme der Unabhängigkeit von irrelevanten Alternativen (IIA) verbunden. Eine teilweise Lockerung dieser Annahme erfolgt beim Nested Logit-Modell, bei dem mehrere Alternativen in Nestern zusammengefasst werden und somit stochastische Nutzenkomponenten teilen. Darüber hinaus bestehen komplexere Modelle mit weniger restriktiven Annahmen.

Eine Überprüfung der IIA-Annahme ist beispielsweise mit dem Hausman-McFadden-Test möglich. Um zu überprüfen, ob signifikante Unterschiede zwischen Nutzenfunktionen unterschiedlicher Gruppen von Individuen oder zwischen unterschiedlichen (restringierten und unrestringierten) Modellen bestehen, kann der (zweistufige) Likelihood-Ratio-Test eingesetzt werden.

4 Verfahren zur Erhebung bei komplexen Multi-Attribut-Entscheidungen

Unabhängig davon, welches Modell zur Auswertung der Discrete Choice Experimente verwendet wird, ist die Anzahl der Attribute zur Charakterisierung der Alternativen bei der Datenerhebung mit gängigen Wahlexperimenten begrenzt. Die Komplexität dieser Experimente steigt mit der Anzahl zusätzlicher Attribute, welche die Alternativen charakterisieren. Für den Befragten ist es mit einer größeren Anstrengung verbunden, alle Information aufzunehmen, zu vergleichen und darauf aufbauend eine Wahlentscheidung zu treffen.²⁰⁰ Es besteht das Risiko der Informationsüberflutung. Als Folge dessen erhöht sich die Wahrscheinlichkeit, dass ein Experiment vorzeitig durch den Befragten abgebrochen wird, Ergebnisse in Folge von Ermüdung verzerrt sind und/oder vereinfachende Strategien verwendet werden, mit denen ebenfalls eine Ergebnisverzerrung einhergeht.²⁰¹ Eine vereinfachende Strategie könnte etwa darin bestehen, dass ein Befragter nur die wichtigsten Attribute berücksichtigt und weniger wichtige Attribute ignoriert. Dabei ist die resultierende Verzerrung davon abhängig, wie ausgeprägt der Einfluss dieser ignorierten Attribute normalerweise wäre.²⁰² Eine weitere vereinfachende Strategie könnte darin liegen, dass ein Befragter jeweils nur die ersten Attribute in seine Entscheidung einbezieht.

Allerdings besteht eine Gefahr darin, die Anzahl der potenziell beeinflussenden Attribute ad hoc oder vor dem Hintergrund empirischer Studien im Vorfeld künstlich zu begrenzen.²⁰³ Es kann zum einen keine Aussage über den Einfluss der nicht analysierten Attribute getroffen werden. Zum anderen könnten von den Befragten die nicht in den Experimenten berücksichtigten Attribute mit ähnlichen berücksichtigten

²⁰⁰ Vgl. *Molin/Timmermans* (2009), S. 635, *Timmermans/Molin* (2009), S. 561 und *Pullman/Dodson/Moore* (1999), S. 126.

²⁰¹ Vgl. *Timmermans/Molin* (2009), S. 561.

²⁰² Vgl. *Molin/Timmermans* (2009), S. 635 und *Timmermans/Molin* (2009), S. 562.

²⁰³ Vgl. *Louviere* (1984), S. 148 und *Timmermans/Molin* (2009), S. 559.

Attributen vermengt werden, was im Ergebnis zu falschen Schlussfolgerungen führen kann.²⁰⁴

4.1 Allgemeine Verfahren

Es existieren mehrere Verfahren²⁰⁵ für die experimentelle Erhebung komplexer Multi-Attribut-Entscheidungen²⁰⁶, d. h. Entscheidungen mit einer größeren Anzahl von Attributen. Beim sogenannten „Pairwise Trade-Off“-Ansatz gilt es, die Ausprägungskombinationen von jeweils zwei Attributen, die in einer sogenannten Trade-off-Matrix dargestellt sind, in eine Rangfolge zu bringen. Dieser Vorgang wird entsprechend oft mit unterschiedlichen Attributspaaren wiederholt.²⁰⁷

Bei „Self-Explication“-Ansätzen, die zu den kompositionellen Verfahren zählen, werden beispielsweise einzelne Attributsausprägungen zunächst auf einer Rating-Skala bewertet; die Ermittlung der Wichtigkeit der Attribute erfolgt anschließend beispielsweise durch Aufteilung einer Gesamtpunktzahl auf die jeweiligen Attribute.²⁰⁸

Hybride Verfahren kombinieren den „Self-Explication“-Ansatz, bei dem zunächst der Teilnutzen für einzelne Attributsausprägungen individuen-spezifisch bestimmt wird, mit einem dekompositionellen Ansatz, bei dem anschließend ausgewählte Attribute gleichzeitig bewertet werden.²⁰⁹ Ein solches hybrides Verfahren stellt die adaptive Conjoint-Analyse dar. Dabei wird zunächst ein vereinfachter „Self-Explication“-Ansatz verwendet, um die Attribute zu identifizieren, die dem jeweiligen Befragten am wichtigsten sind. Im Anschluss daran soll in Paarvergleichen die bevorzugte Alternative und die Stärke der Vorziehenswürdigkeit auf einer

²⁰⁴ Vgl. *Timmermans/Molin* (2009), S. 559.

²⁰⁵ Vgl. für eine detaillierte Beschreibung der Verfahren sowie deren Vergleich beispielsweise *Green/Srinivasan* (1990), *Green/Krieger/Agarwal* (1993), *Pullman/Dodson/Moore* (1999), *Helm/Steiner* (2008), S. 208–237 und *Voeth* (2000), S. 117–133 sowie die dort zitierte Literatur.

²⁰⁶ Allgemein wird die Bezeichnung „Multi-Attribut-Entscheidungen“ verwendet, um auszudrücken, dass eine Alternative durch mindestens zwei Attribute charakterisiert wird, vgl. *Helm/Steiner* (2008), S. 213. Die Bezeichnung „komplexe Multi-Attribut-Entscheidungen“ bezieht sich in Anlehnung an *Louviere* (1984) im Rahmen dieser Arbeit auf Alternativen, die durch mindestens zehn Attribute beschrieben werden.

²⁰⁷ Vgl. *Johnson* (1974), S. 122 und *Helm/Steiner* (2008), S. 213–214.

²⁰⁸ Vgl. *Green/Srinivasan* (1990), S. 9–10 und *Helm/Steiner* (2008), S. 208–211.

²⁰⁹ Vgl. *Helm/Steiner* (2008), S. 216–217 und *Green/Srinivasan* (1990), S. 10–11.

Rating-Skala angegeben werden, wobei die Alternativen nur durch wenige Attribute beschrieben und die jeweiligen Ausprägungen im Rahmen der Erhebung sequentiell angepasst werden.²¹⁰

Beim Bridging-Conjoint-Verfahren werden in unterschiedlichen Experimenten die Alternativen durch unterschiedliche Subsets von Attributen beschrieben, wobei mindestens ein Attribut (Brückenattribut) in allen Experimenten verwendet wird. Über dieses Brückenattribut werden die Attribute der unterschiedlichen Experimente bei der Datenauswertung miteinander verbunden.²¹¹

Wesentliche Nachteile dieser Verfahren bestehen darin, dass Befragte entweder einzelne Attribute bzw. deren Ausprägungen unabhängig von anderen Attributen bewerten müssen (und somit die Annahme erforderlich ist, dass die Attribute voneinander unabhängig sind), eine unvollständige Charakterisierung der Alternativen als unrealistisch wahrgenommen wird und/oder von den Befragten Annahmen über ausgelassene Attribute getroffen werden müssen.²¹² Darüber hinaus werden derartige Verfahren dahingehend kritisiert, dass eine weitreichende theoretische Fundierung fehle.²¹³

Eine größere Anzahl von Attributen kann bei dem Verfahren der hierarchischen Informationsintegration (HII) sowie dessen Weiterentwicklung, dem Verfahren der integrierten hierarchischen Informationsintegration (HII-I), berücksichtigt werden. Diese Verfahren werden im Folgenden beschrieben.

²¹⁰ Vgl. *Sawtooth Software* (2007), S. 2–9, *Green/Srinivasan* (1990), S. 11 und *Helm/Steiner* (2008), S. 220–222.

²¹¹ Vgl. *Albaum* (1989), S. 486–487, *Helm/Steiner* (2008), S. 229 und *Green/Srinivasan* (1990), S. 11.

²¹² Vgl. *Pullman/Dodson/Moore* (1999), S. 127 und *Louviere/Timmermans* (1990a), S. 220.

²¹³ Vgl. *Oppewal/Louviere/Timmermans* (1994), S. 92 und *Louviere/Flynn/Carson* (2010), S. 58–62.

4.2 Verfahren der hierarchischen Informationsintegration (HII)

Das Verfahren der hierarchischen Informationsintegration (HII) wurde in seiner ursprünglichen Form von *Louviere* (1984) vorgeschlagen. Es basiert auf der psychologischen Theorie der Informationsintegration nach *Anderson* (1981 und 1982).

4.2.1 Theorie der Informationsintegration

Der Theorie der Informationsintegration liegen vier ineinandergreifende Konzepte zugrunde.²¹⁴ Gemäß der „Integration“ hängen Gedanken oder Verhalten von Individuen in der Regel nicht von einem, sondern von mehreren Stimuli²¹⁵ ab. Es ist daher von Interesse, wie die Stimuli integriert, d. h. zusammengefasst werden, um eine Reaktion des Individuums auszulösen. Gegenstand der Bewertung (*Valuation*) ist, wie beobachtbare, sogenannte physische Stimuli in sogenannte psychologische Stimuli umgewandelt werden. Die physischen Stimuli können in Experimenten kontrolliert werden; die Integration erfolgt jedoch für die psychologischen Stimuli. Die Integration der Stimuli folgt oftmals einfachen algebraischen Regeln, mit denen sich das Konzept der „kognitiven Algebra“ beschäftigt. Die kognitive Algebra impliziert, dass die Stimuli numerisch dargestellt werden können. Methoden zur Messung der Werte finden im Konzept der „funktionalen Messung“ Berücksichtigung. Von besonderem Interesse ist dabei die sogenannte Integrationsfunktion, welche den Zusammenhang zwischen den psychologischen Stimuli und der daraus folgenden psychologischen Reaktion beschreibt.²¹⁶

Viele Stimuli weisen eine komplexe Struktur auf. So wie sich ein Text in Abschnitte unterteilen lässt, die aus Sätzen bestehen, die sich wiederum aus Satzteilen zusammensetzen, die Wörter enthalten, etc., können auch Stimuli in unterschiedliche Ebenen unterteilt werden. Zur Erklärung der Zusammenhänge der unterschiedlichen Ebenen können die miteinander verbundenen Konzepte der Bewertung und der Integration beitragen:

²¹⁴ Vgl. *Anderson* (1981), S. 2–3 und *Anderson* (1982), S. 4–5.

²¹⁵ Zur Unterscheidung wird hier bewusst nicht der Begriff Attribut verwendet.

²¹⁶ Vgl. *Anderson* (1982), S. 4.

Stimuli der unteren Ebene werden bewertet und zu einer Einheit der darüber liegenden Ebene integriert. Eine Zusammenfassung über mehrere Ebenen ergibt sich als Sequenz von mehreren Integrationen, wobei die Integration auf jeder Ebenen zu einer Vereinfachung führt. Auf diese Weise werden Gruppen von komplexen Stimuli zu einer gemeinsamen Einheit zusammengefasst. Dieser Prozess wird als „Unitization“ bezeichnet.²¹⁷

4.2.2 Grundannahmen der HII

Bei dem Verfahren der hierarchischen Informationsintegration wird bei komplexen (Wahl-)Entscheidungen eine hierarchische Vereinfachungsstrategie unterstellt.²¹⁸ Diese ermöglicht es einem Individuum, eine größere Anzahl von Attributen zu berücksichtigen.²¹⁹

Dem Verfahren der hierarchischen Informationsintegration liegen folgende Annahmen zugrunde:²²⁰

1. Eine Bewertung oder (Wahl-)Entscheidung wird durch ein Set von Attributen beeinflusst. Diese Attribute lassen sich identifizieren und können quantitativer oder qualitativer Natur sein. Die Gesamtzahl der beeinflussenden Attribute wird mit N bezeichnet.
2. Bei komplexen (Wahl-)Entscheidungen verwenden Individuen eine hierarchische Strategie, bei der sie mehrere Attribute zu Subsets zusammenfassen. Diese Subsets werden (Entscheidungs-)Konstrukte genannt. Es wird angenommen, dass die N beeinflussenden Attribute zu insgesamt K Konstrukten zusammenfasst werden. Die Konstrukte werden mit A_k (mit $k = 1, \dots, K$) bezeichnet. Ein Konstrukt besteht aus einem Subset von $N(k)$ Attributen, welche mit $\{x_{k1}, \dots, x_{kN}(k)\}$ bezeichnet werden. Da jedes Attribut in genau einem Konstrukt enthalten ist, entspricht die Gesamtzahl der Attribute der Summe der Attribute in den einzelnen Konstrukten, d. h. es gilt $N = \sum_{k=1}^K N(k)$. Die Anzahl der Attribute der Konstrukte kann

²¹⁷ Vgl. *Anderson* (1981), S. 8.

²¹⁸ Vgl. *Louviere/Timmermans* (1990c), S. 293 und *Louviere* (1984), S. 148–149.

²¹⁹ Vgl. *Louviere* (1984), S. 148 und *Molin/Timmermans* (2009), S. 636.

²²⁰ Vgl. *Louviere/Timmermans* (1990c), S. 293–294 und *Louviere/Timmermans* (1990b), S. 130–131.

variieren, jedoch wird im Sinne einer leichter verständlichen Notation angenommen, dass alle Konstrukte gleich viele Attribute enthalten. Diese Anzahl wird mit n bezeichnet.

3. Die Ausprägungen der Attribute stiften einen Teilnutzen, den die Individuen in Experimenten in Form von Bewertungen oder (Wahl-) Entscheidungen ausdrücken.
4. Individuen bewerten jedes der Konstrukte A_k durch eine Integration der Teilnutzen, die mit den Attributsausprägungen des jeweiligen Konstrukts verbunden sind. Der Zusammenhang zwischen den Attributsausprägungen und der jeweiligen Konstruktbewertung wird durch die Funktion g_k abgebildet, die je nach Konstrukt variieren kann. Näherungsweise wird sie in einem Modell mit linearen Parametern und Attributen beschrieben.
5. Der Teilnutzen der Attributsausprägungen lässt sich nicht direkt messen. Daher werden in Experimenten die Ausprägungen der Attribute des k -ten Konstrukts A_k systematisch variiert und die Bewertungen der Individuen auf einer geeigneten Skala gemessen. Bei den so gemessenen Bewertungen handelt es sich um eine Transformation der unbekannt, aber wahren Teilnutzen. Die gemessene Konstruktbewertung des k -ten Konstrukts wird mit G_k bezeichnet. Dieser Zusammenhang wird ausgedrückt durch

$$G_k = g_k(A_k) = g_k(x_{k1}, \dots, x_{kn}). \quad (4.1)$$

Die Ausprägungen des k -ten Konstrukts werden als Menge $\{G_k\}$ definiert.

6. Individuen bewerten die gesamte Alternative durch eine Integration der einzelnen Konstruktausprägungen, im Folgenden als zusammenfassende Konstruktwerte bezeichnet.²²¹ Der Zusammenhang zwischen den zusammenfassenden Konstruktwerten G_k und der mit R bezeichneten Gesamtbewertung bzw. der (Wahl-)Entscheidung wird durch die Funktion f abgebildet:

²²¹ Bei der Notation der Konstruktausprägungen unterscheidet *Louviere* (1984) explizit zwischen gemessenen Ausprägungen und den in Experimenten variierten hypothetischen Ausprägungen. Diese Unterscheidung der Notation wird aus Übersichtlichkeitsgründen in dieser Arbeit nicht vorgenommen. Sprachlich erfolgt jedoch dahingehend eine Unterscheidung, dass erstere als Konstruktbewertungen und letztere als (hypothetische,) zusammenfassende Konstruktwerte bezeichnet werden.

$$R = f(G_k) = f(g_k(x_{k1}, \dots, x_{kn})), \quad k = 1, \dots, K. \quad (4.2)$$

Diese Funktion lässt sich näherungsweise in einem Modell mit linearen Parametern und Konstrukten beschreiben.

In Experimenten werden die zusammenfassenden Konstruktwerte G_k aller Konstrukte systematisch variiert und die Bewertungen bzw. (Wahl-) Entscheidungen der Individuen gemessen. Sowohl für f als auch für g_k werden üblicherweise additive Verknüpfungsregeln verwendet, wobei jedoch auch andere Verknüpfungsregeln möglich sind.

4.2.3 Aufbau der Experimente der HII

Basierend auf diesen Annahmen lassen sich HII-Experimente entwickeln. Für die Entwicklung der HII-Experimente ist es erforderlich, die potentiell beeinflussenden Attribute den K Konstrukten zuzuordnen.²²² Dabei kann die Zusammenfassung der Attribute auf Basis von „Logik, empirischer Evidenz oder Theorie“²²³ erfolgen. Zumindest im Bereich der Verkehrsmittelwahl sind die Möglichkeiten einer theoriegestützten Entwicklung von Konstrukten sehr begrenzt. Hinsichtlich empirischer Verfahren führt *Louviere* (1984) die Faktorenanalyse und die Verfahren der multidimensionalen Skalierung als Möglichkeiten zur empirischen Entwicklung von Konstrukten an. Ein solches empirisches Verfahren wurde von *Bos/Molin/van der Heijden* (2002), *Bos et al.* (2003), *Bos* (2004) verwendet (vgl. Kap. 5).

Die HII-Experimente bestehen aus zwei Arten von Experimenten, welche als Subexperiment und Brückenexperiment bezeichnet werden. Der grundsätzliche Aufbau dieser Experimente ist in Abb. 4.1 dargestellt. In einem Subexperiment sind die Attribute eines einzelnen Konstrukts enthalten, deren Ausprägungen systematisch variiert werden. Aufgabe der Befragten ist es, auf einer Skala anzugeben, „wie gut“ das durch die entsprechenden Attributsausprägungen beschriebene Konstrukt ist.²²⁴

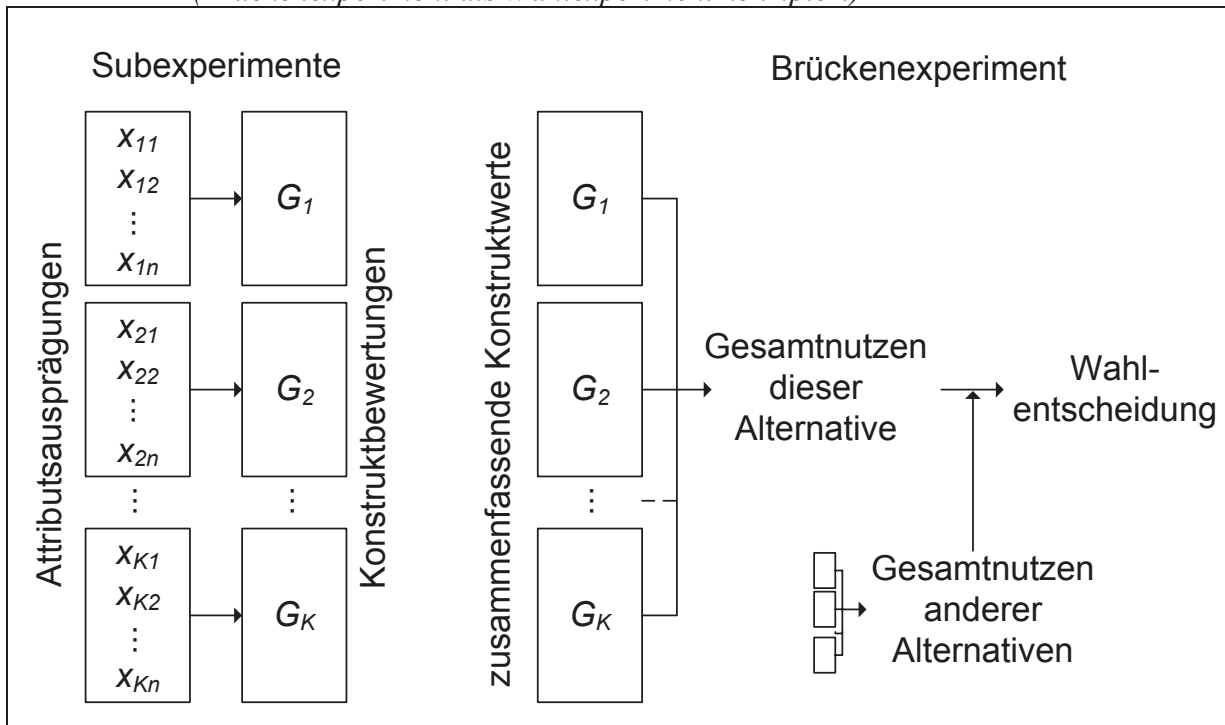
²²² Für eine aggregierte Datenauswertung ist ein einheitliches Design erforderlich, sodass ein jeweiliges Konstrukt stets die gleichen Attribute enthalten muss.

²²³ Vgl. *Louviere/Timmermans* (1990b), S. 132 und *Louviere/Timmermans* (1990c), S. 294.

²²⁴ Vgl. *Louviere* (1984), S. 151.

Üblicherweise werden zu diesem Zweck neun-, zehn- oder elfstufige Rating-Skalen eingesetzt.²²⁵ Da in jedem Subexperiment jeweils nur die Attribute eines Konstrukts enthalten sind, ist bei insgesamt K Konstrukten die Entwicklung von K Subexperimenten erforderlich.

Abb. 4.1: Grundsätzlicher Aufbau eines Experiments der HII (Brückenexperiment als Wahlexperiment konzipiert)



Quelle: Eigene Darstellung in Anlehnung an Louviere (1984, S. 149) und Oppewal/Louviere/Timmermans (1994, S. 93).

Im Brückenexperiment sind alle K Konstrukte enthalten, sodass auf diese Weise die Alternative indirekt durch alle Attribute beschrieben wird. Die Konstrukte werden in Form von hypothetischen, zusammenfassenden Konstruktwerten dargestellt, die systematisch variiert werden. Diese beziehen sich auf die in den Subexperimenten verwendete Skala. Zur Vereinfachung werden jedoch üblicherweise nicht alle, sondern nur einzelne Stufen als Ausprägungen der Konstrukte verwendet, beispielsweise die Ausprägungen 3, 5 und 8 auf einer von 0 bis 10 reichenden Skala.²²⁶ In Abhängigkeit der veränderten Konstruktwerte führt der Befragte eine Gesamtbewertung bzw. (Wahl-)Entscheidung durch. In der von

²²⁵ Bei den Ausprägungen kann es sich mitunter um numerische Werte, beispielsweise eine Skala von 1 bis 10, oder um eine entsprechende Anzahl von Plus- und Minuszeichen handeln, welche anschließend numerisch kodiert werden, vgl. beispielsweise van Helvoort-Postulart et al. (2009a), S. 918.

²²⁶ Vgl. Louviere/Gaeth (1987), S. 32.

Louviere (1984) ursprünglich vorgeschlagenen Form erfolgt in dem Brückenexperiment eine Bewertung auf einer (weiteren) Rating-Skala.²²⁷ Allerdings erfordert der Einsatz einer Rating-Skala die Verwendung einer Entscheidungsregel, um einen Zusammenhang zwischen Bewertungen und Wahlentscheidungen herzustellen. Eine solche Entscheidungsregel stellt mitunter die Annahme dar, dass von mehreren Alternativen jeweils die Alternative mit der höchsten Bewertung gewählt wird. Diese rein deterministische Entscheidungsregel ist jedoch nicht theoretisch fundiert.²²⁸

Zur (teilweisen) Umgehung dieser Problematik haben *Louviere/Timmermans* (1990b) das Brückenexperiment als Wahlexperiment konzipiert. Dabei stehen mehrere Alternativen zur Wahl, die sich hinsichtlich der hypothetischen, zusammenfassenden Konstruktwerte unterscheiden. Die Subexperimente wurden nicht durch ein Wahlexperiment ersetzt, da die Verwendung der gemeinsamen (Rating-)Skala erforderlich ist, um einen Bezug zwischen den Konstruktbewertungen in den Subexperimenten und den hypothetischen, zusammenfassenden Konstruktwerten des Brückenexperiments herzustellen.

4.2.4 Analyse der Experimente der HII

Getrennt für die K Subexperimente wird der Einfluss der Attributsausprägungen auf die Konstruktbewertungen analysiert. Diese Auswertung erfolgt üblicherweise anhand einer multiplen linearen Regressionsanalyse, aber der Einsatz von nicht linearen Modellen wie Ordered Logit oder Ordered Probit ist möglich.²²⁹ Beispielsweise verwendete *Hensher* (1991, S. 306–308) ein Ordered Probit-Modell zur Auswertung der Ratingdaten. Dieses bietet im Gegensatz zur multiplen linearen Regressionsanalyse den Vorteil, dass nur die Rangreihung der Bewertungen berücksichtigt wird und daher im Gegensatz zu linearen Verfahren nicht die Annahme getroffen werden muss, dass der Abstand zwischen den einzelnen Ausprägungen äquidistant ist, d. h. ein mindestens intervallskaliertes Skalen-

²²⁷ Vgl. *Louviere/Gaeth* (1987), S. 151.

²²⁸ Vgl. *Louviere/Woodworth* (1983), S. 351.

²²⁹ Vgl. *Timmermans/Molin* (2009), S. 639.

niveau vorliegt.²³⁰ Neben der Überprüfung, ob die Attributsausprägungen einen signifikanten Einfluss auf die Konstruktbewertung ausüben, lässt sich überprüfen, ob zwischen den Attributsausprägungen eines Konstrukts Interaktionen bestehen.

Sofern das Brückenexperiment als Rating-Experiment konzipiert ist, können zur Auswertung die gleichen Verfahren wie bei den Subexperimenten eingesetzt werden. Wurde das Brückenexperiment hingegen als Discrete Choice Experiment gestaltet, sind entsprechende Verfahren zur Auswertung der diskreten Wahlentscheidungen einzusetzen, beispielsweise ein multinomiales Logit-Modell unter Verwendung der Maximum Likelihood-Schätzung. In den entsprechenden Nutzenfunktionen sind die zusammenfassenden Konstruktwerte mit zugehörigen Nutzenparametern enthalten. In Anlehnung an (3.14) beträgt die Wahlwahrscheinlichkeit, dass Alternative i gewählt wird

$$P(i) = \frac{e^{\mu\gamma'G_i}}{\sum_{j=1}^J e^{\mu\gamma'G_j}} \quad (4.3)$$

wobei es sich bei γ um den Vektor der Nutzenparameter und bei G um den Vektor der zusammenfassenden Konstruktwerte der Alternativen i bzw. j handelt.

Mit geeigneten Schätzverfahren lässt sich der Einfluss der einzelnen zusammenfassenden Konstruktwerte auf die Gesamtbewertung bzw. die Wahlentscheidung bestimmen. Zudem können Interaktionen zwischen Konstruktwerten analysiert werden.

Die unterschiedlichen Subexperimente und das Brückenexperiment lassen sich zu einem Gesamtmodell vereinen. Diese Zusammenfassung der Modelle erfolgt, indem die zusammenfassenden Konstruktwerte durch die jeweiligen Modelle der Subexperimente substituiert werden. Im Ergebnis besteht das Gesamtmodell aus den geschätzten Parametern des Brückenexperiments, den geschätzten Parametern der Subexperimente sowie den Attributsausprägungen. Um diese Substitutionen vornehmen zu können,

²³⁰ Die Annahme, dass Ratingdaten intervallskaliert sind, ist auch in Disziplinen wie der Psychologie weit verbreitet. Diese Annahme hat jedoch hauptsächlich pragmatische Gründe, vgl. *Sedlmeier/Renkewitz* (2008), S. 66.

ist jedoch die Annahme erforderlich, dass die Residuen der einzelnen Modelle jeweils einen Erwartungswert von 0 aufweisen und nicht mit Residuen der anderen Modelle korreliert sind.²³¹

4.2.5 Verfahrensvarianten der HII

Neben der Unterscheidung, ob das Brückenexperiment als Rating- oder Wahlexperiment konzipiert ist, kann dahingehend unterschieden werden, ob eine Alternative im Brückenexperiment ausschließlich durch zusammenfassende Konstruktwerte oder darüber hinaus auch durch Attributsausprägungen charakterisiert ist.²³² Diese Variante wird mitunter in Verkehrsmittelwahlexperimenten eingesetzt, um beispielsweise Zeit- und Preisattribute im Brückenexperiment zu berücksichtigen, wie beispielsweise in den von *Bos et al.* (2004, S. 1676–1677), *Bos* (2004, S. 70–71) und *Norojono/Young* (2003, S. 203–205) beschriebenen Studien. Im Gegensatz zu den zusammenfassenden Konstruktwerten, die sich auf die in den Subexperimenten verwendete Skala beziehen, beziehen sich die Attribute nicht auf diese Skala. Stattdessen werden gängige Einheiten verwendet, beispielsweise für das Preisattribut ein Geldbetrag in der Landeswährung und für das Zeitattribut eine Dauer in Minuten. Die in dem Brückenexperiment berücksichtigten Attribute lassen sich jedoch auch als Konstrukte interpretieren, die nicht auf der Skala der Subexperimente gemessen werden.²³³ So lässt sich beispielsweise ein Zeitattribut in einem Verkehrsmittelwahlexperiment als Konstrukt interpretieren, welches sich aus einzelnen Zeiten wie Reisezeit im Verkehrsmittel, Wartezeit, Umsteigezeit, etc. zusammensetzt.²³⁴

Eine andere Variante besteht in dem von *Ülengin* (1998), *Ülengin/Ülengin/Güvenç* (2001) und *Ülengin/Ülengin Füsün/Gücedinç* (2002) verwendeten Verfahren. Dabei werden in den Subexperimenten zwar Kombinationen von Attributsausprägungen anhand einer Rating-Skala

²³¹ Vgl. *Louviere* (1984), S. 150, *Louviere/Timmermans* (1990c), S. 295 und *Louviere/Timmermans* (1990b), S. 132.

²³² Vgl. *Molin/Timmermans* (2009), S. 641.

²³³ Vgl. *Bos et al.* (2004), S. 1676–1677.

²³⁴ Ebenso könnte ein „weiches“ Qualitätsmerkmal wie z. B. der Komfort eines Sitzplatzes als Konstrukt interpretiert werden, dessen einzelne Attribute wie z. B. Breite des Sitzplatzes (in cm), Sitzreihenabstand (in cm), verwendetes Material etc. direkt messbar sind.

bewertet, allerdings werden die Konstrukte im Brückenexperiment nicht als zusammenfassende Konstruktwerte berücksichtigt, deren Ausprägungen sich auf die Skala der Subexperimente beziehen. Stattdessen handelt es sich bei den zusammenfassenden Konstruktwerten um qualitative Eigenschaften ausgewählter Attribute, sodass kein Zusammenhang zwischen den Subexperimenten und dem Brückenexperiment hergestellt werden kann.²³⁵ *Perrey* (1998) gestaltet die Subexperimente als Ranking-Experimente²³⁶, sodass ebenso kein direkter Zusammenhang zwischen den Subexperimenten und dem Brückenexperiment hergestellt werden kann. Der Bezug der Attributsausprägungen zur Gesamtbewertung bzw. (Wahl-)Entscheidung ist daher nicht gegeben.

4.2.6 Testverfahren

Das Verfahren der hierarchischen Informationsintegration basiert wie in Kap. 4.2.2 beschrieben auf der Annahme einer hierarchischen Struktur. Es existiert jedoch kein spezielles Testverfahren, mit dem die hierarchische Struktur überprüft werden kann. Allerdings bestehen allgemeinere Ansätze zur Überprüfung der Validität des HII-Experiments.

Einer der Ansätze zur Analyse der Validität der HII-Experimente liegt darin, einen Vergleich mit einem sogenannten „Full-Profile“ (FP) Experiment, in dem die jeweilige Alternative durch alle Attribute gleichzeitig beschreiben wird, vorzunehmen. *Louviere* (1984) hat einen solchen Vergleich anhand der geschätzten Parameter durchgeführt und festgestellt, dass zwischen den Parametern des FP-Experiments und denen des Gesamtmodells des HII-Experiments ein linearer Zusammenhang besteht, die Experimente sich somit nur hinsichtlich eines Skalierungsfaktors unterscheiden.²³⁷ Bei diesem Anwendungsfall war die Anzahl der Attribute mit sechs vergleichsweise gering.

Louviere/Gaeth (1987) haben ein Rating-Experiment mit insgesamt 22 Attributen als HII-Experiment (mit vier Konstrukten) sowie als FP-Experiment durchgeführt. Zum Vergleich der beiden Verfahren wurden die Parameter des HII-Gesamtmodells verwendet, um für die jeweiligen

²³⁵ Vgl. *Ülengin* (1998), S. 205 und *Molin/Timmermans* (2009), S. 645.

²³⁶ Vgl. *Perrey* (1998), S. 174–182 und *Perrey* (2000), S. 92–99.

²³⁷ Vgl. *Louviere* (1984), S. 153–154.

Kombinationen von Attributsausprägungen des FP-Experiments Gesamtbewertungen vorherzusagen. Diese vorhergesagten Gesamtbewertungen wurden mit dem Mittelwert der jeweiligen tatsächlichen Bewertungen des FP-Experiments verglichen. Zwischen den Bewertungen wurde ein monotoner, aber nicht linearer Zusammenhang festgestellt. Aus dem signifikanten Rangkorrelationskoeffizienten (Kendalls Tau = 0,66) wurde gefolgert, dass eine Verwendung des heuristischen Verfahrens, demzufolge eine Alternative mit der höchsten Bewertung/dem höchsten Nutzen gewählt wird, gleiche Ergebnisse für das HII- und das FP-Verfahren erzielen würde.²³⁸

Oppewal/Klabbers (2003) haben einen Vergleich zwischen zwei HII-Experimenten, in denen insgesamt 13 Attribute auf unterschiedliche Weise zu Konstrukten zusammenfasst wurden, und einem FP-Experiment vorgenommen. Diese Experimente wurden als Rating-Experimente konzipiert. Zusätzlich wurde ein FP-Experiment als Wahlexperiment gestaltet. Die Rating-Experimente wurden anhand der Modellgüte (adjustiertes R^2) verglichen. Mit 0,67 bzw. 0,55 wiesen die beiden unterschiedlichen Brückenexperimente eine höhere Modellgüte als das FP-Experiment mit 0,38 auf. Die Modellgüte der Subexperimente lag zwischen 0,21 und 0,54 sowie zwischen 0,24 und 0,52 und war somit teilweise besser und teilweise schlechter als die des FP-Experiments. Eine Gesamtaussage hinsichtlich der Modellgüte war somit nicht möglich. Mithilfe der Parameter der unterschiedlichen Modelle wurde für die in den FP-Wahlexperimenten verwendeten Kombinationen von Attributsausprägungen Wahlentscheidungen nach dem heuristischen Verfahren (vgl. 3.5.3) vorhergesagt. Für den Vergleich mit den tatsächlichen Wahlentscheidungen wurden Trefferquoten bestimmt. Die beiden HII-Experimente wiesen mit 81 % bzw. 82 % höhere Trefferquoten als das FP-Experiment mit 77 % auf, allerdings war dieser Unterschied nicht signifikant. Ein weiterer Vergleich wurde anhand der (Nutzen-)Parameter durchgeführt. Sowohl die Parameter des FP-Rating-Experiments (Korrelationskoeffizient $r = 0,90$) als auch die der HII-Gesamtmodelle ($r = 0,95$ bzw. $r = 0,85$) waren mit den Nutzenparametern des FP-Wahlexperiments signifikant korreliert.²³⁹

²³⁸ Vgl. *Louviere/Gaeth* (1987), S. 40.

²³⁹ Vgl. *Oppewal/Klabbers* (2003), S. 303.

Die empirische Aussagekraft derartiger Vergleiche zwischen HII- und FP-Experimenten kann jedoch infrage gestellt werden.²⁴⁰ Mit einem HII-Experiment kann ein anderer Informationsverarbeitungsprozess als mit einem FP-Experiment verbunden sein. Dabei ist unbekannt, welches der Verfahren den „wahren“ Prozess am besten abbildet. Mitunter wird bei der Bestimmung einer Trefferquote unterstellt, dass im FP-Experimente die „wahre“ Gesamtbewertung bzw. Wahlentscheidung erhoben wird, die es im HII-Experiment zu replizieren gilt.²⁴¹ Es kann aber vermutet werden, dass mit zunehmender Anzahl von Attributen ein hierarchisches Verfahren geeigneter ist.²⁴²

Eine Überprüfung der externen Validität von HII-Experimenten führten *Louviere/Timmermans* (1992) durch. Als Anwendungsfall diente die Wahl von Naherholungsgebieten, welche durch insgesamt 19 Attribute beschrieben wurden, die durch vier Konstrukte zusammengefasst wurden. In den Brückenexperimenten wurden sowohl Bewertungen als auch Wahlentscheidungen (mit „no-option“) abgefragt.²⁴³ Zudem wurden Real-daten erhoben. Dazu wählten die Befragten aus einer Liste mit 40 Naherholungsgebieten die fünf von ihnen zuletzt besuchten Gebiete aus. Die tatsächlichen Wahlanteile wurden in Form der relativen Häufigkeit der Nennungen der Naherholungsgebiete differenziert nach Wohngebieten der Befragten bestimmt. Diese Anteile galt es zu replizieren. Für die 40 real existierenden Alternativen wurden die jeweiligen Ausprägungen bestimmt.²⁴⁴ Anhand der Parameter des HII-Gesamtmodells wurden die Wahlanteile der Alternativen vorhergesagt und unterschiedliche Übereinstimmungsmaße berechnet. Es wurde ein deutlicher signifikanter Zusammenhang zwischen den vorhergesagten und den tatsächlichen

²⁴⁰ Vgl. *Louviere* (1984), S. 154 und *Louviere/Gaeth* (1987), S. 43.

²⁴¹ Darüber hinaus besteht das Problem, dass bei Rating-Experimenten Bewertungen in Wahlentscheidungen umgewandelt werden müssen.

²⁴² Vgl. *Molin/Timmermans* (2009), S. 649.

²⁴³ Die von *Louviere/Timmermans* (1992) beschriebenen Ergebnisse beziehen sich jedoch nur auf die Daten des Wahlexperiments, da dadurch direkt aggregierte Wahlanteile bestimmt werden können.

²⁴⁴ Idealerweise hätte erhoben werden müssen, wie die Befragten die Alternativen in Bezug auf die Attributsausprägungen wahrnahmen. Aufgrund der Vielzahl der Attribute und Alternativen konnten diese Daten jedoch nicht erhoben werden und wurden daher nach möglichst objektiven Kriterien bestimmt. Mit dieser Vorgehensweise können leichte Verzerrungen verbunden sein, vgl. *Louviere/Timmermans* (1992), S. 189.

Wahlanteilen (Korrelationskoeffizient $r = 0,68$) festgestellt. Zudem bestand zwischen den Wahlanteilen ein proportionaler Zusammenhang.²⁴⁵

Molin/Timmermans (2003) überprüften in einem Anwendungsfall, ob Individuen in Bezug auf Wohnstätten hierarchisch strukturierte Präferenzen besitzen. Zu diesem Zweck wurden Realdaten erhoben. Dabei beschrieben Befragte ihre aktuelle Wohnstätte in Bezug auf zehn Attribute und bewerteten sie in Bezug auf zwei Konstrukte sowie hinsichtlich der Gesamtzufriedenheit auf einer Rating-Skala. Mithilfe eines Strukturgleichungsmodells wurden die Hypothesen überprüft, dass die Konstrukt-bewertung ausschließlich von den Ausprägungen der zugehörigen Attribute, sowie, dass die Gesamtbewertung ausschließlich von den Konstrukt-bewertungen abhängen. Diese Hypothesen konnten nicht abgelehnt werden, sodass bei diesem konkreten Anwendungsfall von einer hierarchischen Struktur ausgegangen wurde.²⁴⁶ Dieses Verfahren kann im Rahmen einer Vorstudie zur Entwicklung von HII-Experimenten eingesetzt werden. Ein Test der Struktur im HII-Experiment selbst ist hingegen nicht möglich.²⁴⁷

4.2.7 Kritik an dem Verfahren der HII

Trotz einiger erfolgreicher empirischer Ergebnisse weist das Verfahren der hierarchischen Informationsintegration einige Probleme und Grenzen auf:²⁴⁸

- Beim HII-Ansatz liegen getrennte Modelle für Subexperimente und Brückenexperiment vor. Zwischen den variierten Attributsausprägungen und der Gesamtbewertung bzw. der Wahlentscheidung besteht kein direkter Zusammenhang. Ein gemeinsames Modell ist nicht schätzbar, sondern nur durch Substitution der zusammenfassenden Konstruktwerte im Brückenexperiment möglich. Somit liegen anstelle eines Modells mehrere Modelle mit mehreren Kennziffern bezüglich Modellgüte und Validitätstests vor.

²⁴⁵ Vgl. *Louviere/Timmermans* (1992), S. 189.

²⁴⁶ Vgl. *Molin/Timmermans* (2003), S. 46–57.

²⁴⁷ Vgl. *Molin/Timmermans* (2009), S. 650.

²⁴⁸ Vgl. *Oppewal/Louviere/Timmermans* (1994), S. 94 und *Molin/Timmermans* (2009), S. 642–643.

- Es ist nicht bekannt, welche Schwierigkeit für den Befragten mit der Verwendung von hypothetischen, zusammenfassenden Konstruktwerten im eher abstrakten Brückenexperiment verbunden ist. Es kann nicht überprüft werden, ob der Befragte bei seiner Gesamtbewertung bzw. Wahlentscheidung einen hypothetischen, zusammenfassenden Konstruktwert tatsächlich stellvertretend für die entsprechende Konstruktbewertung der zugehörigen Attribute auffasst. Brückenexperimente, die ausschließlich aus numerischen Konstruktwerten bestehenden, könnten zudem einen Befragten zur Bildung von Mittelwerten verleiten.²⁴⁹ Daher bestehen Probleme hinsichtlich der Validität des Brückenexperiments.
- In einem Subexperiment sind nur die Attribute eines einzelnen Konstrukts enthalten, nicht aber andere Konstrukte bzw. deren Attribute. Daher muss der Befragte Annahmen hinsichtlich der Ausprägungen der anderen Konstrukte bzw. Attribute treffen. Die getroffenen Annahmen werden durch das Experiment nicht kontrolliert und sind nicht bekannt.
- Es können nur Interaktionen zwischen Attributen eines Konstrukts sowie Interaktionen zwischen unterschiedlichen Konstrukten bestimmt werden. Interaktionen zwischen Attributen unterschiedlicher Konstrukte sowie zwischen Attributen und Konstrukten sind hingegen nicht schätzbar.
- Das Brückenexperiment kann als Wahlexperiment konzipiert werden, nicht aber das Subexperiment. Zur Analyse von Wahlentscheidungen wäre es daher wünschenswert, wenn die Subexperimente als Wahlexperimente gestaltet werden könnten.
- Es kann nur angenommen werden, dass die dem HII-Modell zugrunde liegende Struktur korrekt ist, sodass die unterschiedlichen Modelle zu einem Gesamtmodell vereint werden können. Diese Annahme kann jedoch nicht getestet werden.

Fast alle dieser Probleme und Grenzen können mit dem von *Oppewal/Louviere/Timmermans* (1994) vorgeschlagenen Ansatz der integrierten hierarchischen Informationsintegration überwunden werden.

²⁴⁹ Diese Problematik besteht nicht oder nur in abgeschwächter Form, wenn, wie oben beschrieben, Attribute oder Konstrukte im Brückenexperiment enthalten sind, deren Ausprägungen nicht auf der gemeinsamen Skala gemessen werden.

4.3 Verfahren der integrierten hierarchischen Informationsintegration (HII-I)

Bei dem von *Oppewal/Louviere/Timmermans* (1994) vorgeschlagenen Ansatz der integrierten hierarchischen Informationsintegration (HII-I) handelt es sich um eine Weiterentwicklung des Ansatzes der hierarchischen Informationsintegration (HII). Somit bleiben die wesentlichen Annahmen des HII-Ansatzes (vgl. 4.2.2) bestehen. Die Weiterentwicklung besteht darin, dass in einem Subexperiment nicht nur die Attribute eines Konstrukts enthalten sind, sondern die verbleibenden anderen *K-I* Konstrukte explizit berücksichtigt werden. Dadurch müssen Befragte keine (nicht kontrollierbaren) Annahmen über die fehlenden Konstrukte treffen.

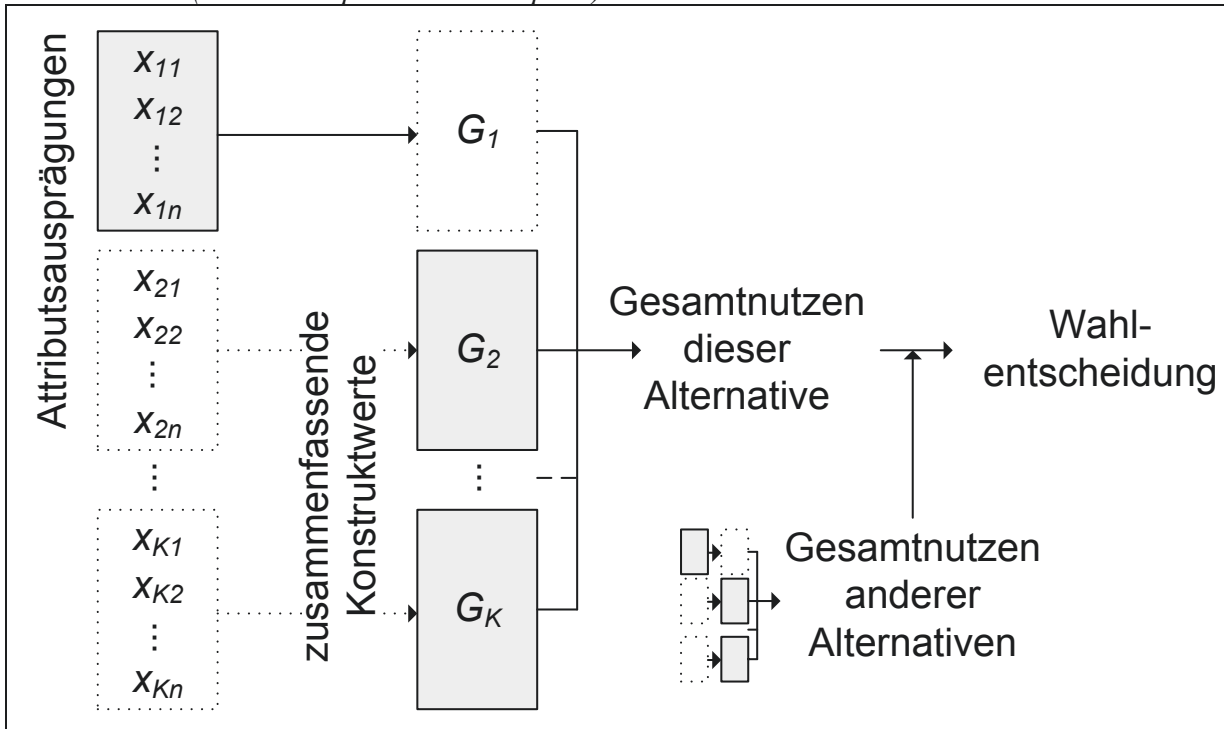
4.3.1 Aufbau der Experimente der HII-I

In einem Subexperiment werden jeweils die Attribute eines Konstrukts sowie zusammenfassende Konstruktwerte der *K-I* anderen Konstrukte variiert. Somit ergeben sich *K* unterschiedliche Subexperimente. Der grundsätzliche Aufbau eines Subexperiments der HII-I ist in Abb. 4.2 dargestellt.

Die Subexperimente können entweder als Rating- oder als Wahlexperimente konzipiert werden. Sofern die Subexperimente als Wahlexperimente gestaltet werden, werden diese auch als integrierte Wahlexperimente (*Integrated Choice Experiments*) bezeichnet. Die Berücksichtigung der anderen Konstrukte dient dazu, die jeweiligen Alternativen vollständig durch alle Attribute zu charakterisieren, entweder direkt durch die Attributsausprägungen oder indirekt durch die zusammenfassenden Konstruktwerte. Infolgedessen dürfte die jeweilige Gesamtbewertung bzw. Wahlentscheidung nicht dadurch beeinflusst werden, welches der *K* Subexperimente verwendet wird. Zudem wird im Gegensatz zum Verfahren der HII ein direkter Zusammenhang zwischen Attributsausprägungen und Gesamtbewertung bzw. Wahlentscheidung hergestellt.

Durch die Berücksichtigung der verbleibenden Konstrukte erhöht sich die Komplexität der Subexperimente. Eine gesonderte Durchführung eines

Abb. 4.2: Grundsätzlicher Aufbau eines Experiments der HII-I (als Wahlexperiment konzipiert)



Quelle: Eigene Darstellung in Anlehnung an Oppewal/Louviere/Timmermans (1994, S. 93). In diesem Subexperiment wird die Alternative durch die grau hinterlegten Attribute und Konstrukte beschrieben.

Brückenexperiments entfällt jedoch, da die Gesamtbewertung bzw. Wahlentscheidung bereits im Rahmen der Subexperimente erfolgt. Zum einen können für die einzelnen Subexperimente eigene Modelle geschätzt werden. Zum anderen können die Modelle der Subexperimente, die (zumindest theoretisch) den gleichen Bewertungs- bzw. Wahlentscheidungsprozess abbilden, zu einem Gesamtmodell vereint werden, in dem alle Attribute und Konstrukte enthalten sind.

Auch bei dem Verfahren der integrierten hierarchischen Informationsintegration ist eine Verfahrensvariante dergestalt möglich, dass ein oder mehrere Attribute, die nicht Teil eines Konstrukts sind, in allen Subexperimenten verwendet werden.

4.3.2 Analyse der Experimente der HII-I

Sofern die Subexperimente als Rating-Experiment konzipiert sind, kann die Datenauswertung beispielsweise anhand einer multiplen linearen Regressionsanalyse oder unter Verwendung von nicht linearen Modellen

wie Ordered Logit oder Ordered Probit erfolgen. Als Wahlexperimente gestaltete Subexperimente können mit den entsprechenden Modellen zur Auswertung von Discrete Choice Daten analysiert werden. Im einfachsten Fall wird dazu ein multinomiales Logit-Modell eingesetzt. Da die jeweiligen Wahlalternativen sowohl durch Ausprägungen von Attributen als auch von Konstrukten beschrieben werden, sind in den jeweiligen Nutzenfunktionen sowohl die Ausprägungen der Attribute als auch der Konstrukte mit den jeweiligen Nutzenparametern enthalten. Die Wahlwahrscheinlichkeit der Alternative i ergibt sich somit als Kombination von (3.14) und (4.3):

$$P(i) = \frac{e^{\mu(\beta'x_i + \gamma'G_i)}}{\sum_{j=1}^J e^{\mu(\beta'x_j + \gamma'G_j)}} \quad (4.4)$$

Im Rahmen der Parameterschätzung ist es beim Verfahren der HII-I im Gegensatz zum Verfahren der HII möglich, Interaktionen zwischen Attributen und Konstrukten zu bestimmen. Die Bestimmung von Interaktionen zwischen Attributen unterschiedlicher Konstrukte ist allerdings nicht möglich. Die dem Modell der HII-I zugrunde liegende hierarchische Struktur kann anhand eines Testverfahrens überprüft werden.

Separate Rating-Experimente, in denen ausschließlich Kombinationen von Attributsausprägungen eines Konstrukts zu bewerten sind, sind grundsätzlich nicht erforderlich. Allerdings kann diese Art von Experimenten dazu eingesetzt werden, den Befragten mit den in den Konstrukten enthaltenen Attributen und deren Ausprägungen sowie der verwendeten Bewertungsskala vertraut zu machen. Zum einen können derartige Rating-Experimente den eigentlichen Subexperimenten vorangestellt werden. Insbesondere bei Computer gestützten und/oder persönlichen Interviews ist dadurch eine Trennung von Rating- und Subexperimenten relativ leicht möglich. Dadurch kann sichergestellt werden, dass die Bewertung der Attributsausprägungen unabhängig von den hypothetischen, zusammenfassenden Konstruktwerten im Subexperiment erfolgt.²⁵⁰ Zum anderen kann das Rating-Experiment (zumindest zur Bewertung eines Konstrukts) in das eigentliche Subexperiment integriert werden. In diesem Fall werden

²⁵⁰ Vgl. *Molin/Timmermans* (2009), S. 650.

die Befragten gebeten, zunächst die Kombinationen der Attributsausprägungen des „zerlegten“ Konstrukts auf der entsprechenden Skala zu bewerten und anschließend zusätzlich die Ausprägungen der anderen Konstrukte bei der Gesamtbewertung bzw. Wahlentscheidung zu berücksichtigen. Jedoch kann bei den in das Subexperiment integrierten Rating-Experimenten nicht sichergestellt werden, dass die Konstruktbewertung unabhängig von den zusammenfassenden Konstruktwerten der anderen Konstrukte erfolgt.

4.3.3 Testverfahren

Oppewal/Louviere/Timmermans (1994) schlagen zur Überprüfung der Validität eines Modells der integrierten hierarchischen Informationsintegration vier Testverfahren vor. Dabei erfordern die ersten drei Verfahren die Durchführung des Rating-Experiments, in dem für Kombinationen von Attributsausprägungen das zugehörige Konstrukt zu bewerten ist.

Mithilfe des ersten Verfahrens soll überprüft werden, ob die Bewertung eines Konstrukts abhängig von den zugehörigen Attributen (bzw. deren Ausprägungen) sowie unabhängig von den anderen Konstrukten ist. Zu diesem Zweck wird – getrennt nach den K unterschiedlichen Rating-Experimenten – ein lineares Regressionsmodell²⁵¹ geschätzt (Konstruktbewertung als abhängige Variable, Ausprägungen der jeweiligen Attribute sowie zusammenfassende Konstruktwerte als unabhängige Variablen). Für eine korrekte Beschreibung eines Konstrukts ist es erforderlich, dass für alle Attribute des Konstrukts signifikante Parameter geschätzt werden, die Parameter der anderen Konstrukte hingegen nicht signifikant sind. Sofern das Rating-Experiment nicht in das Subexperiment integriert wurde, kann das Testverfahren dahingehend geändert werden, dass lediglich überprüft wird, ob alle Attribute einen signifikanten Einfluss auf die Konstruktbewertung ausüben.

Beim zweiten Testverfahren werden die geschätzten Modelle der Subexperimente (zur Abbildung der Gesamtbewertung bzw. der Wahlentscheidung) mit den geschätzten Modellen der Rating-Experimente (zur Abbil-

²⁵¹ Anstelle des linearen Regressionsmodells kann in diesem und den folgenden Testverfahren auch ein geeignetes nicht lineares Modell verwendet werden.

derung der Konstruktbewertung) in Bezug auf signifikante Parameter miteinander verglichen. Es ist erforderlich, dass ein Attribut, das einen signifikanten Einfluss auf die Gesamtbewertung bzw. die Wahlentscheidung ausübt, ebenfalls einen signifikanten Einfluss auf die Konstruktbewertung ausübt. Demzufolge soll ein Parameter, der im Modell des Subexperiments signifikant ist, im Modell des Rating-Experiments ebenfalls signifikant sein. Sofern ein Parameter im Modell des Subexperiments signifikant ist, aber nicht im Modell des Rating-Experiments, kann gefolgert werden, dass das Attribut durch das Konstrukt nicht gut wiedergegeben wird.

Das dritte Testverfahren ermöglicht eine Überprüfung der Prozessgleichheit. Dabei werden die Ergebnisse des k -ten Regressionsmodells eingesetzt, um für die Attributsausprägungen, die sich auf das k -te Konstrukt beziehen, eine Konstruktbewertung vorherzusagen. Im entsprechenden Subexperiment-Modell werden die Attribute durch den vorhergesagten Konstruktwert ersetzt. Folglich besteht das Modell ausschließlich aus vorhergesagten und hypothetischen, zusammenfassenden Konstruktwerten sowie deren Nutzenparametern. *Oppewal/Louviere/Timmermans* (1994) schlagen vor, mithilfe eines modifizierten Chow-Tests²⁵² zu überprüfen, ob eine signifikante Veränderung erfolgt, wenn nicht für die einzelnen Attribute, sondern für den vorhergesagten Konstruktwert Parameter geschätzt werden. Der Chow-Test lässt sich jedoch nur verwenden, wenn das Subexperiment mit einer multiplen linearen Regressionsanalyse ausgewertet wird.²⁵³

Das vierte Testverfahren dient zur Überprüfung der Prozessgleichheit zwischen den unterschiedlichen Subexperimenten. Im Gegensatz zu den anderen Testverfahren basiert dieses Testverfahren ausschließlich auf den Daten der Subexperimente. Eine Durchführung von (separaten oder integrierten) Rating-Experimenten, in denen Kombinationen von Attributsausprägungen eines Konstrukts zu bewerten sind, ist nicht erforderlich. Da jedes der K unterschiedlichen Subexperimente den gleichen Prozess hinsichtlich der Gesamtbewertung bzw. der Wahlentscheidung abbilden soll, dürfen sich die Parameter, die sich auf ein bestimmtes

²⁵² Vgl. zum modifizierten Chow-Test beispielsweise *Kmenta* (1997), S. 420–422.

²⁵³ Vgl. *Kmenta* (1997), S. 421.

Konstrukt beziehen, nicht zwischen den Subexperimenten (mit Ausnahme von Stichprobenfehler und Fehlervarianz)²⁵⁴ unterscheiden.²⁵⁵ *Oppewal/Louviere/Timmermans* (1994) schlagen diesbezüglich die Verwendung eines modifizierten Chow-Tests vor. Sofern die Subexperimente jedoch als Wahlexperimente gestaltet sind, wird anstelle des modifizierten Chow-Tests ein zweistufiger Likelihood-Ratio-Test verwendet.²⁵⁶ In der ersten Stufe des Testverfahrens besteht die Restriktion darin, dass die Nutzenparameter der Konstrukte (sowie der Attribute, die nicht Teil eines Konstrukts sind und in allen Subexperimenten verwendet werden) auf Gleichheit restringiert werden. In der zweiten Stufe werden neben den gemeinsamen Nutzenparametern auch die Skalierungsparameter auf Gleichheit restringiert (vgl. Kap. 3.5.2).

Sofern zwischen den Subexperimenten Prozessgleichheit besteht, kann für diese Modelle ein Gesamtmodell geschätzt werden, welches alle Attribute und Konstrukte enthält. Besteht jedoch keine Prozessgleichheit, ist das jeweilige Subexperiment-Modell kontextabhängig, d. h. die Gesamtbewertung bzw. die Wahlentscheidung unterscheidet sich je nach Subexperiment. Die Kontextabhängigkeit muss bei der Interpretation der Ergebnisse berücksichtigt werden; jedes einzelne Subexperiment kann aber wie ein normales (nicht hierarchisches) Wahlexperiment angesehen werden.²⁵⁷

4.4 Anwendungsbeispiele der Verfahren der HII und der HII-I

Das Verfahren der hierarchischen Informationsintegration (und dessen Varianten) sowie dessen Weiterentwicklung, das Verfahren der integrierten hierarchischen Informationsintegration, wurde bereits in unterschiedlichen Bereichen angewendet.

²⁵⁴ Vgl. *Oppewal/Louviere/Timmermans* (1994), S. 96.

²⁵⁵ Jedes der K Konstrukte wird in $K-1$ Subexperimenten in Form eines hypothetischen, zusammenfassenden Konstruktswerts verwendet. Damit die Gleichheit der sich auf ein Konstrukt beziehenden Parameter überprüft werden kann, muss dieses Konstrukt in mindestens zwei Subexperimenten verwendet werden, vgl. *van de Vyvere/Oppewal/Timmermans* (1998), S. 260.

²⁵⁶ Vgl. *van de Vyvere/Oppewal/Timmermans* (1998), S. 260–261.

²⁵⁷ Vgl. *Oppewal/Louviere/Timmermans* (1994), S. 96.

Anwendungen außerhalb des Verkehrsbereichs analysieren mitunter Präferenzen bzw. Wahlentscheidungen hinsichtlich Telekommunikation²⁵⁸, Einkaufsstätten²⁵⁹, Wohnstätten²⁶⁰, Bürostandorten²⁶¹, Naherholungsgebieten²⁶², Bankunternehmen²⁶³ und Gesundheitsversorgung.²⁶⁴

Bei den Anwendungen im Verkehrsbereich wurde bislang ausschließlich das Verfahren der HII oder dessen Variante, nicht aber das Verfahren der HII-I eingesetzt. Die Analysen beziehen sich sowohl auf den Personen- als auch auf den Güterverkehr.

Hensher (1991) analysierte in einer Verfahrensvariante der hierarchischen Informationsintegration Präferenzen von 3696 Nutzern und 100 Nicht-Nutzern des privaten Busverkehrs. Die Daten wurden in einer schriftlichen Befragung erhoben. Die Subexperimente enthielten vier, drei, fünf bzw. drei Attribute, die sich auf die Konstrukte Fahrt-, Warte-, Fahrzeug- bzw. Informationsqualität bezogen. In dem als Rating-Experiment konzipierten Brückenexperiment waren neben zusammenfassenden Konstruktwerten auch die Attribute Fahrpreis und intramodale Umsteigenotwendigkeit enthalten.²⁶⁵

Chiang/Lu/Chang (2003) untersuchten den Einfluss von Qualitätsattributen auf die intermodale Verkehrsmittelwahl für Reisen mit zwei unterschiedlichen Distanzen (Luftlinie zwischen den Städten: ca. 270 bzw. 135 km). In persönlichen Befragungen mit Wartenden am Bahnhof, an der Bushaltestelle und am Flughafen wurden insgesamt 735 gültige Interviews durchgeführt, die sich gleichmäßig auf die beiden Distanzen aufteilten. Die Subexperimente beinhalteten sechs, vier, drei bzw. drei Attribute der Konstrukte Dienstleistungsqualität, Qualität der Weiterreise am Ziel(bus)bahnhof/Zielflughafen, Informationsqualität bzw. Qualität der Umgebung. Das Brückenexperiment wurde als bimodales Wahlexperiment gestaltet. Dabei standen für die längere Distanz ein Hochgeschwin-

²⁵⁸ Vgl. *Louviere* (1984) und *Ramírez-Hurtado* (2010).

²⁵⁹ Vgl. *Louviere/Gaeth* (1987), *Oppewal/Louviere/Timmermans* (1994) und *Oppewal et al.* (1997).

²⁶⁰ Vgl. *Louviere/Timmermans* (1990b), *van de Vyvere/Oppewal/Timmermans* (1998) und *Oppewal/Klabbers* (2003).

²⁶¹ Vgl. *Argioli et al.* (2008).

²⁶² Vgl. *Louviere/Timmermans* (1990c) und *Louviere/Timmermans* (1992).

²⁶³ Vgl. *Oppewal/Vriens* (2000).

²⁶⁴ Vgl. *van Helvoort-Postulart et al.* (2009a) und *van Helvoort-Postulart et al.* (2009b).

²⁶⁵ Vgl. *Hensher* (1991), S. 302.

digkeitszug und ein Flugzeug, für die kürzere Distanz ein Zug und ein Bus als Wahlalternativen zur Verfügung.²⁶⁶

Bos (2004), *Bos et al.* (2004) und *Bos/Molin* (2006) analysierten den Einfluss der Attribute von P+R (*Park and Ride*) Anlagen auf die (multimodale) Verkehrsmittelwahl. In postalischen und internetbasierten Befragungen wurden gültige Daten von insgesamt 805 Pkw-Nutzern erhoben. Die Subexperimente bestanden aus sieben bzw. vier Attributen der Konstrukte Qualität der P+R Anlage bzw. Qualität des daran angeschlossenen ÖPNV. Das Brückenexperiment wurde als trimodales Wahlexperiment durchgeführt. Die Wahl erfolgte vor dem Hintergrund von Kontextvariablen²⁶⁷ zwischen einer multimodalen Alternative (mit dem Pkw zur P+R Anlage, anschließend öffentliches Verkehrsmittel), einer Pkw-Alternative und einer ÖPNV-Alternative. Die Ausprägungen der ÖPNV-Alternative wurden nicht verändert. Die multimodale Alternative wurde durch zusammenfassende Werte der beiden Konstrukte beschrieben. Darüber wurden sowohl die multimodale als auch die Pkw-Alternative durch relative Zeit- und Preisattribute im Vergleich zur ÖPNV beschrieben.²⁶⁸

In Anlehnung an diese Anwendung untersuchten *Molin/van Gelder* (2008) ebenfalls den Einfluss auf die (multimodale) Verkehrsmittelwahl. In einer internetbasierten Befragung wurden gültige Daten von insgesamt 384 Pkw- und ÖPNV-Nutzern erhoben. Neben den Konstrukten Qualität der P+R Anlage mit sieben Attributen und Qualität des angeschlossenen ÖPNV mit vier Attributen bezog sich ein weiteres Konstrukt, das sechs Attribute umfasste, auf das Angebot des „hoch-qualitativen öffentlichen Verkehrs“. Damit ist eine (hypothetische) Magnetschwebebahn gemeint, welche mehrere Städte ringförmig miteinander verbindet, und in Verbindungen mit anderen Verkehrsmitteln genutzt wird. Für diese drei Konstrukte wurden Subexperimente entwickelt. Das Brückenexperiment wurde als Wahlexperiment mit einer multimodalen Alternative, einer

²⁶⁶ Vgl. *Chiang/Lu/Chang* (2003), S. 719–723.

²⁶⁷ Anhand der Kontextvariablen wurde eine hypothetische Rahmensituation wie beispielsweise Wetterbedingungen, Gepäckmitnahme und Tageszeit beschrieben. Die Analyse zeigte, dass diese Rahmenbedingungen nur einen vergleichsweise geringen Einfluss auf die Verkehrsmittelwahl ausüben, vgl. *Bos et al.* (2004), S. 1682–1683.

²⁶⁸ Vgl. *Bos* (2004), S. 70–72.

Pkw-Alternative und einer ÖPNV-Alternative gestaltet. Die multimodale Alternative und die Pkw-Alternative wurden durch relative Zeit- und Preisattribute im Vergleich zur ÖPNV-Alternative, deren Attribute selbst nicht verändert wurden, charakterisiert. Die multimodale Alternative wurde darüber hinaus durch zusammenfassende Konstruktwerte der drei Konstrukte beschrieben.²⁶⁹

Norojono/Young (2003) analysierten den Einfluss von Qualitätsattributen auf die intermodale Verkehrsmittelwahl im Güterverkehr. In Befragungen²⁷⁰ wurden gültige Daten von insgesamt 186 Verladern erhoben. Die Subexperimente enthielten vier bzw. drei Attribute, die sich auf die Konstrukte Qualität bzw. Flexibilität bezogen. In dem Brückenexperiment wurden die Alternativen kleine Lkw, große Lkw und Bahn durch zusammenfassende Konstruktwerte sowie durch die beiden Attribute Transportkosten und Lieferzeit beschrieben. Dieses Experiment wurde als Ranking-Experiment durchgeführt.²⁷¹

Eine Übersicht der Anwendungen der Verfahren der HII bzw. der HII-I zeigt Tab. 4.1. Aus dieser Übersicht wird ersichtlich, dass bislang nur wenige Anwendungen des integrierten Verfahrens durchgeführt wurden. Entsprechend gibt es nur wenige Anwendungen, in denen die hierarchische Struktur des Modells getestet wurde. *Oppewal/Louviere/Timmermans* (1994) überprüften für eine Anwendung der HII-I die Gleichheit der Nutzenparameter zwischen den Subexperimenten und stellten keine empirische Evidenz für die Gleichheit fest. Für eine andere Anwendung der HII-I lehnten *van de Vyvere/Oppewal/Timmermans* (1998) die Hypothese der Gleichheit der Nutzenparameter nicht ab; es bestanden jedoch Unterschiede zwischen den Skalierungsparametern der Subexperimente.

Darüber hinaus wurde das Verfahren der HII bzw. der HII-I eingesetzt, um Präferenzen/Wahlentscheidungen von Gruppen von Individuen (z. B. Familien) in Bezug auf Wohnstätten zu analysieren.²⁷² Es bestehen dahingehend Abweichungen, dass sich beim Brückenexperiment bzw.

²⁶⁹ Vgl. *Molin/van Gelder* (2008), S. 107–108.

²⁷⁰ Die Art der Befragung geht aus *Norojono/Young* (2003) nicht hervor.

²⁷¹ Vgl. *Norojono/Young* (2003), S. 203–206.

²⁷² Zu derartigen Anwendungen vgl. *Timmermans et al.* (1992), *Molin* (1999), *Molin/Oppewal/Timmermans* (2000) und *Molin/Oppewal/Timmermans* (2001).

Tab. 4.1: Anwendungen des Verfahrens der HII und der HII-I

| Quelle | Anwendungsgebiet und Datengrundlage | Konstrukte (Anzahl Attribute) |
|--|---|--|
| Verfahren der HII, Rating-Experiment | | |
| <i>Louviere</i> (1984) | <ul style="list-style-type: none"> • Telekommunikation • 104 Studenten | <ul style="list-style-type: none"> • Kosten (3) • Service (3) |
| <i>Louviere/Gaeth</i> (1987) | <ul style="list-style-type: none"> • Einkaufsstätten • 62 Studenten | <ul style="list-style-type: none"> • Preis (4) • Qualität (3) • Auswahl (4) • Annehmlichkeiten (11) |
| <i>Hensher</i> (1991) | <ul style="list-style-type: none"> • Personenverkehr (Intramodale Verkehrsmittelwahl) • 3696 Nutzer und 100 Nicht-Nutzer des privaten Busverkehrs | <ul style="list-style-type: none"> • Fahrtqualität (4) • Wartequalität (3) • Fahrzeugqualität (5) • Informationsqualität (3) • Fahrpreis (1) • Intramodale Umsteigenotwendigkeit (1) |
| <i>Bos</i> (2004), <i>Bos et al.</i> (2004), <i>Bos/Molin</i> (2006) | <ul style="list-style-type: none"> • Personenverkehr (Multimodale Verkehrsmittelwahl) • 805 Pkw-Nutzer | <ul style="list-style-type: none"> • Qualität der P+R Anlage (7) • Qualität des angeschlossenen ÖPNV (4) • Zeit (1) • Preis (1) |
| <i>Molin/van Gelder</i> (2008) | <ul style="list-style-type: none"> • Personenverkehr (Multimodale Verkehrsmittelwahl) • 384 Pkw- und ÖPNV-Nutzer | <ul style="list-style-type: none"> • Qualität der P+R Anlage (7) • Qualität des angeschlossenen ÖPNV (4) • Angebot des „hochqualitativen öffentlichen Verkehrs“ (6) • Zeit (1) • Preis (1) |
| <i>Argioli et al.</i> (2008) | <ul style="list-style-type: none"> • Bürostandorte • 404 Geschäftsleiter | <ul style="list-style-type: none"> • Erreichbarkeit des Standorts (5) • Erscheinungsbild des Gebäudes (1) • Nutzung der Innenräume (1) • Mietkosten (1) • Parkplatzverfügbarkeit (1) |

Fortsetzung nächste Seite

Fortsetzung Tab. 4.1

| Quelle | Anwendungsgebiet und Datengrundlage | Konstrukte (Anzahl Attribute) |
|--|--|--|
| Verfahren der HII, Rating- und Wahlexperiment | | |
| <i>Louviere/Timmermans</i> (1990c), <i>Louviere/Timmermans</i> (1992) | <ul style="list-style-type: none"> • Naherholungsgebiete • 364 Einwohner | <ul style="list-style-type: none"> • Umgebung und Erreichbarkeit (5) • Einrichtungen und Aktivitäten (7) • Pflege (4) • Soziales Umfeld (3) |
| Verfahren der HII, Wahlexperiment | | |
| <i>Louviere/Timmermans</i> (1990b) | <ul style="list-style-type: none"> • Wohnstätten • 315 Neueinwohner | <ul style="list-style-type: none"> • Haus (6) • Wohnumfeld (6) • Lage (6) • Soziale und wirtschaftliche Bindungen (4) |
| <i>Chiang/Lu/Chang</i> (2003) | <ul style="list-style-type: none"> • Personenverkehr (Intermodale Verkehrsmittelwahl) • 735 Verkehrsteilnehmer | <ul style="list-style-type: none"> • Dienstleistungsqualität (6) • Qualität der Weiterreise am Ziel(bus)bahnhof/ Zielflughafen (4) • Informationsqualität (3) • Qualität der Umgebung (3) |
| <i>Norojono/Young</i> (2003) | <ul style="list-style-type: none"> • Güterverkehr (Intermodale Verkehrsmittelwahl) • 186 Verloader | <ul style="list-style-type: none"> • Qualität (4) • Flexibilität (3) • Transportkosten (1) • Lieferzeit (1) |
| Verfahren der HII, Wahlexperiment | | |
| <i>Oppewal/Klabbers</i> (2003) | <ul style="list-style-type: none"> • Wohnstätten • 70 Studenten | <ul style="list-style-type: none"> • Größe der Wohnräume (3) • Größe der Schlafräume (3) • Raumanordnung (4) • Sonnenlicht (3) <p>bzw.</p> <ul style="list-style-type: none"> • Wohnzimmer (3) • Küche (3) • Flur und zweites Schlafzimmer (3) • Badezimmer und größtes Schlafzimmer (4) |

Fortsetzung nächste Seite

Fortsetzung Tab. 4.1

| Quelle | Anwendungsgebiet und Datengrundlage | Konstrukte (Anzahl Attribute) |
|---|--|---|
| Verfahren der HII-I, Rating-Experiment | | |
| <i>Oppewal/Vriens</i> (2000) | <ul style="list-style-type: none"> • Banken • 200 Bankkunden | <ul style="list-style-type: none"> • Erreichbarkeit der Bank (7) • Kompetenz der Angestellten (5) • Genauigkeit und Freundlichkeit (9) • Materielle Werte (7) |
| <i>Ramírez-Hurtado</i> (2010) | <ul style="list-style-type: none"> • Mobiltelefone • 110 Studenten | <ul style="list-style-type: none"> • Größe und Preis (3) • Batterieleistung (2) • Anrufeigenschaften (3) • Entertainment (5) • Schnittstellen (3) |
| Verfahren der HII-I, Wahlexperiment | | |
| <i>Oppewal/Louviere/Timmermans</i> (1994), <i>Oppewal et al.</i> (1997) | <ul style="list-style-type: none"> • Einkaufsstätten • 396 bzw. 405 Haushalte (Befragung der Person, die für den Haushalt die meisten Einkäufe durchführt) | <ul style="list-style-type: none"> • Auswahl der Geschäfte, die Kleidung und Schuhe anbieten (9) • Auswahl der Geschäfte, die Lebensmittel und verpackte Ware anbieten (9) • Erscheinungsbild und Gestaltung des Einkaufszentrums (10) • Lage und Erreichbarkeit des Einkaufszentrums (5) |
| <i>van de Vyvere/Oppewal/Timmermans</i> (1998) | <ul style="list-style-type: none"> • Wohnstätten • 120 Neueinwohner | <ul style="list-style-type: none"> • Haus (5) • Wohnumgebung (3) • Lage (6) |
| <i>van Helvoort-Postulart et al.</i> (2009a), <i>van Helvoort-Postulart et al.</i> (2009b) | <ul style="list-style-type: none"> • Ambulante Krankenhausbetreuung • 156 Krankenhausangestellte | <ul style="list-style-type: none"> • Organisation (5) • Kooperationspartner (5) • Patientenorientierte Pflege (5) • Leitfaden (1) • Zeitliche Investition (1) |

Quelle: In Anlehnung an *Molin/Timmermans* (2009, S. 646).

beim integrierten Subexperiment die zusammenfassenden Konstruktwerte auf (hypothetische) Bewertungen der unterschiedlichen Gruppenmitglieder beziehen. Trotz wesentlicher methodischer Übereinstimmungen basiert dieses Verfahren nicht auf den Annahmen, die dem Verfahren der HII bzw. HII-I zugrunde liegen.²⁷³

4.5 Zusammenfassung

Bei der Datenerhebung mit gängigen Wahlexperimenten ist die Anzahl der Attribute zur Charakterisierung der Alternativen begrenzt. Zur Datenerhebung bei sogenannten komplexen Multi-Attribut-Entscheidungen (mehr als zehn Attribute) kann beispielsweise das Verfahren der hierarchischen Informationsintegration (HII) bzw. dessen Weiterentwicklung, das Verfahren der integrierten hierarchischen Informationsintegration (HII-I), eingesetzt werden. Beide Verfahren basieren auf der Annahme, dass Individuen bei komplexen (Wahl-)Entscheidungen eine hierarchische Strategie verwenden, bei der sie ähnliche Attribute zu sogenannten (Entscheidungs-)Konstrukten zusammenfassen und bewerten. Für die Gestaltung von Experimenten der HII und der HII-I ist eine derartige Zusammenfassung von Attributen zu Konstrukten a priori erforderlich.

Bei dem Verfahren der HII werden zunächst sogenannte Subexperimente durchgeführt, in denen Kombinationen von Attributsausprägungen, die sich auf ein Konstrukt beziehen, auf einer geeigneten Skala bewertet werden. In einem sogenannten Brückenexperiment erfolgt anschließend eine Gesamtbeurteilung (Rating- oder Wahlexperiment) für hypothetische, zusammenfassende Konstruktwerte der unterschiedlichen Konstrukte. Dieses Verfahren wird im Wesentlichen dahingehend kritisiert, dass kein direkter Zusammenhang zwischen den Attributsausprägungen und dem Gesamturteil besteht und die dem Modell zugrunde liegende hierarchische Struktur nicht getestet werden kann.

Bei dem Verfahren der HII-I werden die Subexperimente dahingehend geändert, dass nicht nur die Attributsausprägungen eines Konstrukts enthalten sind, sondern auch zusammenfassende Konstruktwerte der anderen

²⁷³ Vgl. *Molin/Timmermans* (2009).

Konstrukte. Dadurch wird jede Alternative durch alle Attribute direkt oder indirekt beschrieben. Ein gesondertes Brückenexperiment erfolgt daher nicht mehr. Die Subexperimente können als Rating- oder Wahlexperimente gestaltet werden. Bei diesem Verfahren ist eine Überprüfung der hierarchischen Struktur mit einem entsprechenden Testverfahren möglich.

Bislang wurden nur wenige Anwendungen des Verfahrens der HII-I durchgeführt. Entsprechend gibt es nur wenige Anwendungen, in denen die hierarchische Struktur des Modells getestet wurde.

5 Verfahren zur empirischen Entwicklung von Konstrukten

Die Verfahren der HII und der HII-I erfordern, dass Attribute a priori zu Konstrukten zusammengefasst werden. Derartige Konstrukte sollten Attribute enthalten, welche Befragte als ähnlich wahrnehmen. Dabei ist es im Sinne eines einheitlichen Befragungsdesigns notwendig, dass für alle Befragten die gleichen Konstrukte verwendet werden.

Zur Entwicklung von Konstrukten können empirische Verfahren eingesetzt werden, welche dazu dienen, Strukturen abzubilden, wie beispielsweise Verfahren der multidimensionalen Skalierung oder der hierarchischen Clusteranalyse. Gegenstand dieser Verfahren ist die Analyse von Proximitäten, d. h. Nähe-, Abstands-, Ähnlichkeits- oder Unähnlichkeitsdaten, welche entweder erhoben oder aus anderen Daten abgeleitet werden.²⁷⁴

Im Folgenden werden zunächst wesentliche Verfahren zur Erhebung von Ähnlichkeitsdaten vorgestellt und ein für diese Arbeit relevantes Verfahren ausgewählt. Im Anschluss daran werden vor dem Hintergrund des Skalenniveaus der Daten Verfahren zur Auswertung und eine mögliche Überprüfung resultierender Ergebnisse aufgezeigt.

5.1 Verfahren zur Erhebung von Ähnlichkeitsdaten

Verfahren zur Erhebung von Ähnlichkeiten können dahingehend unterschieden werden, wie viele Objekte²⁷⁵ ein Befragter gleichzeitig und wie viele Objekte er insgesamt bewertet sowie welche Schwierigkeit damit verbunden ist.²⁷⁶ Bei allen Verfahren werden Z Objekte vorgegeben und die Aufgabe des Befragten liegt darin, Objektpaare in Bezug auf die von

²⁷⁴ Vgl. *Borg* (2010), S. 391.

²⁷⁵ Zur Abgrenzung wird der Begriff Objekt anstelle von Attribut verwendet. Bei den Objekten kann es sich ganz allgemein beispielsweise um Begriffe oder Abbildungen handeln.

²⁷⁶ Vgl. *Bijmolt/Wedel* (1995), S. 364.

ihm wahrgenommene Ähnlichkeit zu beurteilen, in eine Rangfolge zu bringen oder Gruppen mit ähnlichen Objekten zu bilden.

Beim Ratingverfahren²⁷⁷ sind die $Z(Z-1)/2$ Objektpaare jeweils einzeln auf einer Ähnlichkeits- oder Unähnlichkeitsskala zu bewerten. Als Skala wird üblicherweise eine 7- oder 9-stufige bipolare Skala (z. B. „sehr ähnlich“ bis „sehr unähnlich“) verwendet.

Bei der Methode der Rangreihung²⁷⁸ liegt die Aufgabe des Befragten darin, alle $Z(Z-1)/2$ Objektpaare gemäß der empfundenen Ähnlichkeit so zu ordnen, dass eine aufsteigende bzw. absteigende Rangfolge der Ähnlichkeiten entsteht. Zur Vereinfachung kann eine grobe Gruppeneinteilung mit ähnlichen und unähnlichen Paaren durchgeführt werden und anschließend so lange Untergruppen gebildet werden, bis letztendlich eine Rangfolge vorliegt.

Bei der Ankerpunktmethode²⁷⁹ sind alle $Z-1$ Objekte nach ihrer Ähnlichkeit zu einem Vergleichsobjekt (Ankerpunkt) in eine Rangfolge zu bringen. Jedes Objekt dient dabei genau einmal als Ankerpunkt, sodass insgesamt $Z(Z-1)$ Rangwerte gebildet werden. Da die Rangfolge von dem jeweiligen Ankerpunkt abhängig ist, handelt es sich bei den erhobenen Daten um bedingte (konditionale) Daten. Da das Objekt z_1 sich in Bezug auf den Ankerpunkt z_2 nicht notwendigerweise auf dem gleichen Rang befindet wie das Objekt z_2 in Bezug auf den Ankerpunkt z_1 , führt diese Methode in der Regel zu einer asymmetrischen Datenmatrix.²⁸⁰

Ein weiteres konditionales Verfahren stellt der Tripelvergleich²⁸¹ dar. Dabei sind aus jeweils drei gegebenen Objekten das ähnlichste und das unähnlichste Objektpaar zu benennen. Da die Anzahl der Dreiergruppen mit $Z(Z-1)(Z-2)/6$ Kombinationsmöglichkeiten sehr groß ist, wird oft ein Design (z. B. ein „Balanced Incomplete Block Design“) eingesetzt, um die Anzahl der zu bewertenden Dreiergruppen zu reduzieren.

²⁷⁷ Vgl. *Backhaus/Erichson/Weiber* (2011), S. 225.

²⁷⁸ Vgl. *Green/Carbone/Smith* (1989), S. 37–38, *Bijmolt/Wedel* (1995), S. 364 und *Backhaus/Erichson/Weiber* (2011), S. 223.

²⁷⁹ Vgl. *Bijmolt/Wedel* (1995), S. 364, *Neidell* (1972), S. 335 und *Backhaus/Erichson/Weiber* (2011), S. 224.

²⁸⁰ Sofern für die Auswertung eine symmetrische Matrix erforderlich ist, ist eine Datentransformation durchzuführen.

²⁸¹ Vgl. *Bijmolt/Wedel* (1995), S. 364 und *Neidell* (1972), S. 335.

Die genannten Verfahren haben gemeinsam, dass die Anzahl der durchzuführenden Beurteilungen mit der Anzahl der Objekte überproportional zunimmt. Sofern zur Datenauswertung eine Aggregation erforderlich ist, erfolgt diese in der Regel durch Bildung von Mittelwerten oder Medianen. Allerdings weisen die erhobenen Daten ein ordinales Skalenniveau auf, sodass aus mathematischer Sicht nur die Bildung von Medianen, nicht aber von Mittelwerten zulässig ist.

Beim Sorting-Verfahren²⁸² sind die Z Objekte in Bezug auf ein Sortierungskriterium (wie beispielsweise „allgemeine Ähnlichkeit“) so in Gruppen zu sortieren, dass sich die Objekte innerhalb einer Gruppe ähnlich sind, aber von den Objekten anderer Gruppen unterscheiden. Beim Sorting-Verfahren können grundlegend mehrere Varianten anhand folgender Kriterien²⁸³, die miteinander kombiniert werden, unterschieden werden.²⁸⁴

- die Anzahl der Gruppen, in die die Objekte sortiert werden, ist frei wählbar oder vorgegeben (free vs. fixed),
- die Objekte werden nur genau einer Gruppe zugeordnet oder können mehreren Gruppen zugeordnet werden (disjoint/non-overlapping vs. overlapping),
- das Sortierungskriterium ist frei wählbar oder vorgegeben,
- die Gruppennamen sind frei wählbar oder vorgegeben (unlabelled vs. labelled),
- der Sortiervorgang erfolgt einmalig oder mehrfach zur Bildung von Gruppen nach unterschiedlichen Kriterien (single vs. multiple).

Bei dem sogenannten Disjoint-Free-Sorting, auch verkürzt als Free-Sorting bezeichnet, ist jedes Objekt in genau eine Gruppe zu sortieren, wobei sowohl die Anzahl der Gruppen als auch deren Namen von dem Befragten selbst zu bestimmen ist.

²⁸² Vgl. *Coxon* (1999), S. 1–4, *Miller* (1969), S. 169–170, *Rosenberg/Park Kim* (1975), S. 489 und *Spencer* (2009), S. 4–5.

²⁸³ Weitere Varianten erlauben beispielsweise die Bildung einer Hierarchie oder geben neben der Anzahl der Gruppen auch die Anzahl der Objekte innerhalb einer jeweiligen Gruppe vor (Q-sort).

²⁸⁴ Vgl. *Coxon* (1999), S. 1-4, 18-27.

Im Gegensatz zu den vorher beschriebenen Verfahren kann beim Verfahren des Free-Sortings nicht die Stärke der Ähnlichkeit zwischen Objektpaaren durch den Befragten ausgedrückt werden; stattdessen werden Ähnlichkeiten von Objektpaaren nur dichotom ausgedrückt (entweder „ähnlich“ oder „unähnlich“), je nachdem ob die Objekte sich in einer gemeinsamen Gruppe befinden oder nicht.

Bijmolt/Wedel (1995) haben in einer Studie das Verfahren des Free-Sortings, das Ratingverfahren, die Ankerpunktmethode und den Triplevergleich²⁸⁵ miteinander verglichen. Dabei wurden die Befragten zufällig einem der Verfahren zugeordnet und es wurden Ähnlichkeitsdaten für zwölf Objekte erhoben, wobei die Dauer der Datenerhebung gemessen wurde. Im Anschluss wurden die Befragten gebeten, das Verfahren auf einer Skala von 0 bis 100 hinsichtlich der Ermüdung und Langeweile infolge der Datenerhebung, der Menge der zur Verfügung stehenden Informationen sowie die Verständlichkeit der Aufgabe zu bewerten. Das Sorting-Verfahren war das schnellste Verfahren (3,2 min), gefolgt von dem Ratingverfahren (7,8 min), dem Triplevergleich (18,7 min) und der Ankerpunktmethode (19,3 min). Des Weiteren handelt es sich bei dem Sorting-Verfahren um das Verfahren, das die geringste Ermüdung und Langeweile verursacht (Mittelwerte: 16,2 bzw. 25,8), gefolgt von dem Ratingverfahren (24,1 bzw. 30,9), dem Triplevergleich (31,7 bzw. 42,6) und der Ankerpunktmethode (32,6 bzw. 40,0). Hinsichtlich der Menge der zur Verfügung stehenden Informationen und der Verständlichkeit der Aufgabe wurden keine signifikanten Unterschiede festgestellt. Zusammenfassend wurde empfohlen, die Ankerpunktmethode oder den Triplevergleich einzusetzen, wenn die Anzahl der Objekte gering ist und eine größtmögliche Menge an Information erwünscht ist. Bei einer großen Anzahl an Objekten ist hingegen das Sorting-Verfahren vorzuziehen, sofern keine Auswertung der Daten auf Individualebene erfolgen soll; ansonsten stellt das Ratingverfahren einen geeigneten Kompromiss zwischen der Befragungsdauer und der Menge an erhobenen Informationen dar.

²⁸⁵ Beim Triplevergleich wurde die Anzahl der Dreiergruppen von 220 auf 80 reduziert.

In der Studie von *Bijmolt/Wedel* (1995) wurden lediglich zwölf Objekte auf ihre Ähnlichkeit untersucht. Aufgrund der überproportional steigenden Anzahl von Objektbewertungen beim Ratingverfahren und bei der Ankerpunktmethode ist bei einer größeren Anzahl von Objekten mit noch stärkeren Unterschieden zwischen den Verfahren zu rechnen.

Im Rahmen der vorliegenden Arbeit wird eine größere Anzahl von Objekten verwendet. Zudem sollen die hergeleiteten Konstrukte in III-I Experimenten gleichermaßen für alle Befragten eingesetzt werden sollen, sodass eine differenzierte Auswertung der Ähnlichkeitsdaten auf Individualebene nicht erforderlich ist. Vor diesem Hintergrund erscheint das Verfahren des Free-Sortings zur Erhebung von Ähnlichkeitsdaten geeignet.

Daten, die mit dem Verfahren des Free-Sortings erhoben wurden, lassen sich differenziert für jeden Befragten in Form einer Ähnlichkeits- oder Unähnlichkeitsmatrix²⁸⁶ auf Individualebene ausdrücken. Dabei handelt es sich um eine quadratische Matrix, deren Zeilen- und Spaltenzahl der Anzahl der Objekte entspricht. In der Zelle c_{qr} der Unähnlichkeitsmatrix wird die Häufigkeit des gemeinsamen Nicht-Auftretens der Objekte q und r ²⁸⁷ (d. h. wie oft sich dieses Objektpaar in unterschiedlichen Gruppen befindet) erfasst. Folglich handelt es sich um eine symmetrische Matrix, deren Zellen ausschließlich die Werte 0 und 1 enthalten. Die Erstellung einer aggregierten Unähnlichkeitsmatrix erfolgt durch Summation aller Unähnlichkeitsmatrizen auf Individualebene. Kleine Werte in den Zellen dieser Unähnlichkeitsmatrix deuten auf sehr ähnliche Objektpaare hin (viele Befragte haben die Objekte in die gleiche Gruppe sortiert), große Werte hingegen auf unähnliche Objektpaare. Der kleinste mögliche Wert beträgt 0 und tritt nur dann auf, wenn alle Befragten das betreffende Objektpaar in eine gemeinsame Gruppe sortiert haben. Der größte mögliche Wert entspricht der Anzahl der Befragten und tritt nur dann auf, wenn das jeweilige Objektpaar von allen Befragten in unterschiedliche

²⁸⁶ Für die meisten Statistikprogramme ist zur Datenauswertung eine Unähnlichkeitsmatrix erforderlich, sodass im Folgenden nur auf die Eigenschaften der Unähnlichkeitsmatrix eingegangen wird.

²⁸⁷ Das gemeinsame Auftreten von Objekten wird auch als „Co-Occurrence“ bezeichnet, vgl. *Borg* (2010), S. 391.

Gruppen sortiert wurde.²⁸⁸ Die Daten der aggregierten Unähnlichkeitsmatrix weisen ein ratioskaliertes (vgl. Anhang A) Skalenniveau auf, sodass ein metrisches Verfahren zur Datenauswertung eingesetzt werden kann. Zwar können auch Verfahren für ein niedrigeres Skalenniveau eingesetzt werden, jedoch ist dies mit einem Informationsverlust verbunden.

5.2 Multidimensionale Skalierung

Als Multidimensionale Skalierung (MDS) wird eine Gruppe von Verfahren bezeichnet, bei denen Objekte in Form von Punkten in einem mehrdimensionalen²⁸⁹ Raum so angeordnet werden, dass die Distanzen zwischen diesen Objektpunkten den zugehörigen Proximitäten bestmöglich entsprechen.²⁹⁰ In der Regel werden diese Verfahren explorativ²⁹¹ eingesetzt, um Strukturen in Daten aufzudecken und zu visualisieren. Das Aufdecken von Strukturen ist insbesondere dann von Interesse, wenn diese nicht theoretisch hergeleitet werden können.

Die räumliche Darstellung der Objektpunkte wird als Konfiguration bezeichnet.²⁹² Handelt sich bei den darzustellenden Daten um Unähnlichkeitsdaten, sind in der Konfiguration Objekte, die sich sehr ähnlich sind, nah beieinander und hingegen Objekte, die sich sehr unähnlich sind, weit entfernt voneinander angeordnet.

Zur Veranschaulichung der Verfahren der MDS wird in der Literatur häufig ein Beispiel eingesetzt, bei dem basierend auf Entfernungen zwischen Städten eine Landkarte rekonstruiert wird.²⁹³ Als Datengrundlage dient dabei eine Matrix mit paarweisen Entfernungen zwischen Städten, wie sie beispielsweise in einem Straßenatlas zu finden ist. Eine

²⁸⁸ Vgl. *Coxon* (1999), S. 43.

²⁸⁹ Die Anzahl der Dimensionen, bezeichnet als Dimensionalität, ist dabei (extern) vorzugeben.

²⁹⁰ Vgl. *Borg/Groenen* (2005), S. 3, *Borg* (2010), S. 391, *Cox/Cox* (2001), S. 5–6 und *Kruskal/Wish* (1978), S. 7.

²⁹¹ Der Einsatz einer MDS als konfirmatorisches Verfahren ist möglich, um Annahmen über bestehende Strukturen zu überprüfen, vgl. *Borg/Groenen* (2005), S. 6.

²⁹² Vgl. *Kruskal/Wish* (1978), S. 7.

²⁹³ Vgl. dazu *Kruskal/Wish* (1978), S. 8, *Cox/Cox* (2001), S. 2, *Borg/Groenen* (2005), S. 19–23 und *Backhaus/Erichson/Weiber* (2011), S. 219–222.

ermittelte Konfiguration in einem zweidimensionalen Raum, in diesem Fall die rekonstruierte Landkarte, bildet die relative Lage der Städte adäquat ab.²⁹⁴ Die Daten werden oftmals vergleichend mit zwei unterschiedlichen Verfahren der MDS (metrische und nicht-metrische MDS) ausgewertet. Obwohl dabei beide Verfahren zu annähernd gleichen Konfigurationen führen, sollten daraus keine verallgemeinernden Schlussfolgerungen gezogen und ein Verfahren entsprechend des vorliegenden Skalenniveaus gewählt werden.²⁹⁵

Übereinstimmendes Ziel der Verfahren der MDS ist die Ermittlung einer Konfiguration. Wesentlich ist dabei neben der Wahl einer Abbildungsfunktion und eines Distanzmodells die Wahl eines Kriteriums, welches die Anpassungsgüte der Konfiguration beschreibt. Die Anpassungsgüte gilt es unter Verwendung eines geeigneten Algorithmus zu optimieren. Die Verfahren der MDS unterscheiden sich hinsichtlich dieser Aspekte.²⁹⁶ Die folgenden Ausführungen beziehen sich schwerpunktmäßig auf ein Verfahren der metrischen MDS²⁹⁷, da die Daten der aggregierten Unähnlichkeitsmatrix wie in Anhang A beschrieben ratioskaliertes Skalenniveau aufweisen.

Allgemein werden bei der multidimensionalen Skalierung Proximitäten p_{qr} als Distanzen $d_{qr}(\mathbf{X})$ innerhalb einer Konfiguration \mathbf{X} dargestellt. Bei der MDS für ratioskalierte Daten soll das Verhältnis eines beliebigen Paares von Proximitäten dem Verhältnis des zugehörigen Paares von Distanzen entsprechen.²⁹⁸ Grundsätzlich wird der Zusammenhang zwischen Proximitäten und Distanzen beschrieben durch die Abbildungsfunktion

$$f: p_{qr} \rightarrow d_{qr}(\mathbf{X}). \quad (5.1)$$

²⁹⁴ Zur Erstellung einer nach Norden ausgerichteten Landkarte ist in der Regel eine Spiegelung und Rotation der ermittelten Konfiguration erforderlich.

²⁹⁵ Vgl. *Borg* (2010), S. 399.

²⁹⁶ Vgl. *Solaro* (2012), S. 359–364 und *Borg* (2010), S. 399.

²⁹⁷ Vgl. bezüglich der nicht-metrischen MDS beispielsweise *Kruskal* (1964a), *Kruskal* (1964b), *Kruskal/Wish* (1978), *Borg/Groenen* (2005) und *Hartung/Elpelt* (1995), S. 405–421.

²⁹⁸ Vgl. *Borg/Groenen* (2005), S. 200–201.

Die Distanzen bzw. die Lage der Objektpunkte im m -dimensionalen Raum sind unbekannt und im Rahmen der MDS zu ermitteln. Dabei sind die Distanzen so zu bestimmen, dass sie idealerweise vollständig den – unter Berücksichtigung des Skalenniveaus – transformierten Proximitäten entsprechen:

$$f(p_{qr}) = d_{qr}(\mathbf{X}). \quad (5.2)$$

Bei der MDS für ratioskalierte Daten ist diese Funktion definiert durch

$$f(p_{qr}) = b \cdot p_{qr}, \quad (5.3)$$

mit $b > 0$, sofern Unähnlichkeitsdaten verwendet werden. Transformierte Proximitäten werden auch als Disparitäten bezeichnet.²⁹⁹

Das Distanzmodell gibt an, wie die Distanzen zwischen den Objektpunkten innerhalb der Konfiguration berechnet werden. Die euklidische Distanz bietet den Vorteil der leichten Interpretierbarkeit der Konfiguration, denn die Distanz zwischen zwei Punkten entspricht der Länge der Geraden zwischen diesen Punkten („Luftlinie“). Zudem ist die ermittelte Konfiguration rotationsinvariant, d. h. eine Rotation der Koordinatenachsen führt zu keiner Veränderung der Distanzen. Aus diesem Grund eignet sich die euklidische Distanz insbesondere bei einem explorativ eingesetzten Verfahren der MDS.³⁰⁰

Die euklidische Distanz zwischen zwei Punkten q und r beträgt

$$d_{qr} = \left[\sum_{a=1}^m (x_{qa} - x_{ra})^2 \right]^{\frac{1}{2}}, \quad (5.4)$$

wobei es sich bei x um die Koordinaten der jeweiligen Objektpunkte in einem m -dimensionalen Raum handelt.³⁰¹

Idealerweise entsprechen die Distanzen den transformierten Proximitäten wie in (5.2) vollständig, normalerweise bestehen jedoch Abweichungen,

²⁹⁹ Vgl. Cox/Cox (2001), S. 63, Borg (2010), S. 396 und Borg/Groenen (2005), S. 199–201.

³⁰⁰ Vgl. Borg/Groenen (2005), S. 14. Zu anderen Distanzmaßen vgl. beispielsweise Cox/Cox (2001), S. 10–11.

³⁰¹ Vgl. Borg/Groenen (2005), S. 39 und Cox/Cox (2001), S. 11.

sodass die Gleichung nur näherungsweise gilt. Wie gut die transformierten Proximitäten als Distanzen zwischen Punkten wiedergegeben werden, lässt sich durch den (quadrierten) Repräsentationsfehler e^2 ausdrücken, aufsummiert über alle Objektpaare

$$\rho_R(\mathbf{X}) = \sum_{(q<r)} e^2 = \sum_{(q<r)} [f(p_{qr}) - d_{qr}(\mathbf{X})]^2. \quad (5.5)$$

ρ_R wird als Rohstress bezeichnet und beschreibt die Anpassungsgüte der gesamten MDS-Konfiguration.³⁰²

Da der Rohstress von der durchgeführten Transformation – in diesem Fall von dem Wert des Faktors b – sowie von der Anzahl der Objekte abhängt, wird eine Normierung des Rohstresses vorgenommen.³⁰³

Eine Normierung anhand der Summe der quadrierten Disparitäten ergibt den normierten Rohstress

$$\rho_n(\mathbf{X}) = \sqrt{\frac{\sum_{(q<r)} [f(p_{qr}) - d_{qr}(\mathbf{X})]^2}{\sum_{(q<r)} [f(p_{qr})]^2}} \quad (5.6)$$

Die Normierung anhand der Summe der quadrierten Distanzen ergibt den Kruskal's Stress oder Stress-1

$$\rho_1(\mathbf{X}) = \sqrt{\frac{\sum_{(q<r)} [f(p_{qr}) - d_{qr}(\mathbf{X})]^2}{\sum_{(q<r)} d_{qr}^2(\mathbf{X})}}. \quad (5.7)$$

Es handelt sich dabei um die bekannteste Statistik hinsichtlich der Güte einer Konfiguration der MDS.³⁰⁴ Die Verwendung der Quadratwurzel ermöglicht eine bessere Unterscheidung von ansonsten numerisch recht ähnlichen Werten nahe 0.³⁰⁵ Bei einer perfekten Konfiguration stimmen Disparitäten und Distanzen vollständig überein, der Wert für Stress-1

³⁰² Vgl. *Borg/Groenen* (2005), S. 42.

³⁰³ Vgl. *Litz* (2000), S. 437.

³⁰⁴ Vgl. *Kruskal* (1964a), S. 9 und *Borg/Staufenbiel* (2007), S. 163.

³⁰⁵ Vgl. *Borg* (2010), S. 397.

beträgt somit 0. In der Regel bestehen jedoch Abweichungen, die zu einem Wert von größer 0 für Stress-1 führen.³⁰⁶

Neben dem normalisierten Rohstress und Stress-1 existieren weitere Stressmaße, darunter Stress-2 und S-Stress. Im Gegensatz zu Stress-1 erfolgt bei Stress-2 die Normierung nicht mittels der aufsummierten quadrierten Distanzen, sondern mittels der Summe der quadrierten Abweichungen der Distanzen von der durchschnittlichen Distanz \bar{d} :

$$\rho_2(\mathbf{X}) = \sqrt{\frac{\sum_{(q<r)} [f(p_{qr}) - d_{qr}(\mathbf{X})]^2}{\sum_{(q<r)} [d_{qr}(\mathbf{X}) - \bar{d}]^2}}. \quad (5.8)$$

Bei gleicher Konfiguration ergibt sich aufgrund dieser Normierung für Stress-2 ein größerer Wert als für Stress-1.³⁰⁷

S-Stress unterscheidet sich von diesen beiden Stressmaßen dahingehend, dass im Zähler nicht der Rohstress, sondern die aufsummierten quadrierten Differenzen zwischen den quadrierten Disparitäten und den quadrierten Distanzen stehen:

$$\rho_S(X) = \sum_{(q<r)} \left[\left(f(p_{qr}) \right)^2 - d_{qr}^2 \right]^2. \quad (5.9)$$

Auch bei diesem Stressmaß werden entsprechende Veränderungen (Verwendung von Quadratwurzel und Normierungen) vorgenommen. Die Verwendung von S-Stress liegt überwiegend in den algorithmischen Vorteilen (Stressfunktion immer differenzierbar, garantierte Konvergenz des Algorithmus), die mit diesem Stressmaß verbunden sind.³⁰⁸ Dem gegenüber kann jedoch als Nachteil angesehen werden, dass durch das Quadrieren große Unähnlichkeiten stärker ins Gewicht fallen als kleine.³⁰⁹

³⁰⁶ Der Wert von Stress-1 ist dimensionslos und variiert zwischen 0 und einem im Allgemeinen unbekanntem Wert, der deutlich kleiner als 1 ist, vgl. *Borg* (2000), S. 6, *Borg/Groenen* (2005), S. 274–276 und *Cox/Cox* (2001).

³⁰⁷ Vgl. *Borg/Groenen* (2005), S. 251.

³⁰⁸ Vgl. *Takane/Young/Leeuw* (1977), S. 18 und *Borg/Groenen* (2005), S. 252.

³⁰⁹ Vgl. *Borg/Groenen* (2005), S. 252. Zudem kann die Anpassungsgüte in Bezug auf Abweichungen zwischen einzelnen Paaren von Disparitäten und Distanzen analysiert werden, vgl. *Cox/Cox* (2001), S. 72–73 und *Borg/Staufenbiel* (1997), S. 103–104.

Für eine Bewertung der Anpassungsgüte wird oftmals die von *Kruskal* (1964a, S. 3) vorgeschlagene Klassifizierung für den Wert von Stress-1 herangezogen. Sie unterscheidet zwischen einer „perfekten“ ($\rho_1 = 0$), „exzellenten“ ($\rho_1 = 0,025$), „guten“ ($\rho_1 = 0,05$), „ausreichenden“ ($\rho_1 = 0,01$) und „schlechten“ ($\rho_1 = 0,2$) Anpassungsgüte. Eine solche rein formale und absolute Beurteilung der Anpassungsgüte ist allerdings nicht frei von Kritik. Diese besteht darin, dass der Stresswert von mehreren Faktoren beeinflusst wird und somit vom jeweiligen Kontext abhängig ist.³¹⁰ Mit zunehmender Anzahl von Objekten und damit einhergehender überproportional zunehmender Anzahl von Disparitäten und Distanzen steigt der Stresswert. Dahingegen nimmt der Stresswert mit zunehmender Anzahl von Dimensionen ab, da zusätzliche Dimensionen mehr Freiheiten zur Positionierung der Objektpunkte bieten. Bei einer metrischen MDS liegen mehr Restriktion vor als bei einer ordinalen (nicht-metrischen) MDS, bei der nur Ranginformation wiedergegeben werden, sodass die metrische MDS in der Regel zu einem höheren Stresswert führt. Zudem deutet eine „schlechte“ Anpassungsgüte nicht notwendigerweise auf eine schlechte Konfiguration hin, sondern kann auch aus „Rauschen“ in den Daten resultieren. Eine präzisere Darstellung, d. h. somit eine Darstellung des Rauschens, würde in diesem Fall keinen inhaltlichen Mehrwert bieten.

Zur eigentlichen Ermittlung der Konfiguration wird ein Algorithmus eingesetzt, mit dem die Anpassungsgüte optimiert wird. Dabei unterscheiden sich die Algorithmen im Wesentlichen dahingehend, welches Stressmaß zur Beschreibung der Anpassungsgüte verwendet wird und wie die Optimierung erfolgt. Da für die meisten MDS-Modelle keine analytische Lösung bekannt ist, erfolgt die Suche nach der optimalen Konfiguration iterativ.³¹¹ Ausgehend von einer Startkonfiguration, welche entweder systematisch (beispielsweise mit dem Simplex- und dem Torgerson-Algorithmus)³¹² oder zufällig (einfache oder mehrfache Zufallsstarts) bestimmt wird, werden die Objektpunkte im Raum schrittweise so verschoben, dass der Stresswert kleiner wird. Ein Minimum des Stresswerts ist erreicht, wenn keine weitere Minimierung des Stresswerts durch

³¹⁰ Vgl. *Borg* (2010), S. 398–399, *Borg/Staufenbiel* (2007), S. 179–180, *Borg/Groenen* (2005), S. 47–57 und *Hamerle/Pape* (1996), S. 788.

³¹¹ Vgl. *Borg/Staufenbiel* (2007), S. 167 und *Solaro* (2012), S. 367.

³¹² Vgl. *IBM Corporation* (2011a), S. 774–775 und die dort zitierte Literatur.

eine Verschiebung der Objektpunkte mehr möglich ist. Dabei muss sichergestellt werden, dass es sich bei dem gefundenen Minimum um ein globales Minimum handelt.³¹³

Zwei bekannte MDS-Algorithmen sind in *SPSS* implementiert. Dabei handelt es sich um den *Alscal*-Algorithmus und dessen Nachfolger, den *Proxscal*-Algorithmus. Beim *Alscal*-Algorithmus wird der S-Stress unter Verwendung des Verfahrens der alternierenden kleinsten Quadrate (*Alternating Least Squares*)³¹⁴ minimiert. Dieser Algorithmus garantiert – im Gegensatz zu älteren Algorithmen – eine Konvergenz gegen das Minimum des Stresswertes. Die Verwendung des S-Stresses führt jedoch dazu, dass nicht eine Anpassung an ursprüngliche Werte (transformierte Proximitäten), sondern an abgeleitete Werte (quadrierte transformierte Proximitäten) erfolgt, welche eine stärkere Gewichtung größerer Werte verursacht.³¹⁵ Beim *Proxscal*-Verfahren wird der normalisierte Rohstress unter Verwendung des Subgradientenverfahrens³¹⁶ minimiert. Dadurch bietet der *Proxscal*-Algorithmus den Vorteil, dass eine Konvergenz garantiert ist, jedoch eine Anpassung an ursprüngliche Werte erfolgt. Bei beiden Algorithmen wird die jeweilige Konfiguration iterativ so lange verbessert, bis die Verbesserung des Stresswerts und/oder der Stresswert selbst einen vorgegebenen Schwellenwert unterschreitet.³¹⁷ Die so ermittelte Konfiguration stellt die Lösung der MDS für die entsprechende Anzahl von Dimensionen dar.

Die Anzahl der Dimensionen muss bei der Ermittlung der Konfiguration extern vorgegeben werden und ist – im Gegensatz zu dem oben beschriebenen Städtebeispiel – in der Regel unbekannt. Bei der metrischen MDS lassen sich Konfiguration mit *Z* Objekten grundsätzlich in maximal *Z*-1 Dimensionen bestimmen. Allerdings ist die Anzahl der Dimensionen zum einen technisch³¹⁸ und zum anderen hinsichtlich der Interpretierbarkeit der

³¹³ Vgl. *Borg/Staufenbiel* (2007), S. 167.

³¹⁴ Vgl. zu diesem Verfahren *Takane/Young/Leeuw* (1977), S. 18–34 und *Young/Takane/Lewyckyj* (1978), S. 433–435.

³¹⁵ Vgl. *Hamerle/Pape* (1996), S. 784.

³¹⁶ Vgl. zu diesem Verfahren *Commandeur/Heiser* (1993), S. 11–16 und *IBM Corporation* (2011a), S. 775–776.

³¹⁷ Vgl. *IBM Corporation* (2011c), S. 314 und *Meulman/Heiser* (2011), S. 83.

³¹⁸ Bei dem *Alscal*-Algorithmus können nur Konfigurationen mit bis zu sechs Dimensionen bestimmt werden, vgl. *IBM Corporation* (2011c), S. 314. Bei dem *Proxscal*-Algorithmus liegt eine derartige Einschränkung nicht vor, vgl. *Meulman/Heiser* (2011), S. 81.

Ergebnisse beschränkt, sodass sich die Lösungen in der Regel auf zwei oder drei Dimensionen beschränken.³¹⁹

Die Auswahl der geeigneten Dimensionalität kann vor dem Hintergrund der Verbesserung des Stresswerts bei zunehmender Anzahl von Dimensionen erfolgen.³²⁰ Dazu ist es zunächst erforderlich, mehrere Konfigurationen mit unterschiedlicher Dimensionalität zu bestimmen. In einem sogenannten Scree-Plot werden in Abhängigkeit der Anzahl der Dimensionen auf der Abszissenachse die Stresswerte auf der Ordinatenachse aufgetragen und die Punkte dieser diskreten Werte miteinander verbunden. Die resultierende Kurve verläuft monoton fallend. Ist in der Kurve ein sogenannter „Ellbogen“ (*Elbow*) vorhanden, d. h. ein Punkt, ab dem das Stressmaß deutlich weniger abnimmt als vorher, so ist die zu diesem Punkt gehörende Dimensionalität zu wählen. Begründen lässt sich die Auswahl dadurch, dass bis zu diesem Punkt die Dimensionen zur Abbildung systematischer Strukturen genutzt werden, zusätzliche Dimensionen hingegen keinen weiteren Erklärungsgehalt liefern.³²¹ Ein derartiger Ellbogen ist jedoch nicht immer vorhanden.³²²

Wenn eine geeignete Dimensionalität ausgewählt ist, kann die zugehörige Konfiguration interpretiert werden. Dabei ist insbesondere die räumliche Lage der Objektpunkte zueinander von Bedeutung. Eine inhaltliche Interpretation der Koordinatenachsen ist – selbst nach Rotation und Spiegelung – nicht immer möglich und stellt einen Spezialfall dar. Allgemeiner ist hingegen die Interpretation von sogenannten Regionen. Dabei handelt es sich um nicht-überlappende Teilgebiete, die zusammen den gesamten MDS-Raum abdecken. Wenn die Objekte einer Region sehr nah beieinander liegen und zudem leerer Raum zwischen den Objekten der unterschiedlichen Regionen liegt, so werden die Regionen auch als Cluster bezeichnet.³²³

Im Rahmen der Entwicklung von Konstrukten ist insbesondere die Identifikation solcher Cluster von Interesse. Im Allgemeinen ist jedoch die Interpretation einer Konfiguration nicht immer eindeutig und es können

³¹⁹ Vgl. Borg/Staufenbiel (2007), S. 179.

³²⁰ Vgl. Hamerle/Pape (1996), S. 789 und Wührer (2000), S. 453–454.

³²¹ Vgl. Borg/Groenen (2005), S. 48.

³²² Vgl. Borg/Groenen (2005), S. 52, 71.

³²³ Vgl. Borg/Staufenbiel (2007), S. 170–173 und Borg/Groenen (2005), S. 81.

mehrere miteinander konkurrierende Interpretationen bestehen. Beispielsweise könnte ein etwas weiter entferntes Objekt noch als Teil eines Clusters oder bereits als eigenes Cluster, das aus nur einem Objekt besteht, interpretiert werden.

Grundsätzlich könnte zur Analyse der Unähnlichkeitsdaten anstelle eines Verfahrens der multidimensionalen Skalierung auch ein Verfahren der (hierarchischen) Clusteranalyse³²⁴ eingesetzt werden. Jedoch ist ein Verfahren der MDS gegenüber einer Clusteranalyse zu bevorzugen, wenn bei der Datenerhebung mit dem Sorting-Verfahren mit paradigmatischen Sortierungen zu rechnen ist, d. h. die Befragten unterschiedliche Kriterien bei der Sortierung heranziehen.³²⁵ Dies ist im Rahmen der vorliegenden Arbeit möglicherweise gegeben, sodass ein Verfahren der MDS geeigneter erscheint.

Eine Konfiguration, die auf aggregierten Unähnlichkeitsdaten basiert, bildet möglicherweise die Sorting-Daten auf Individualebene nicht geeignet ab. Ein Vergleich der (interpretierten) Konfiguration mit den Sorting-Daten auf Individualebene kann daher als Indikator der internen Validität angesehen werden. Dieser kann auch herangezogen werden, um zwischen konkurrierenden Interpretationen auszuwählen.

Es existieren Verfahren zum Vergleich von Lösungen der MDS. Diese basieren jedoch auf den geometrischen Eigenschaften der Konfigurationen³²⁶ und sind daher zum Vergleich der (interpretierten) Konfiguration mit den Sorting-Daten auf Individualebene ungeeignet.

Bei den Sorting-Daten auf Individualebene handelt es sich um Partitionen³²⁷ im mathematischen Sinne. Ebenso kann eine Konfiguration mit mehreren Clustern bzw. klar voneinander trennbaren Regionen

³²⁴ Vgl. zu den Verfahren der (hierarchischen) Clusteranalyse beispielsweise *Kaufmann/Pape* (1996), *Büschken/Thaden* (2000), *Bortz* (2005), S. 571–580 und *Bacher* (1996) sowie die dort zitierte Literatur.

³²⁵ Vgl. *Miller* (1969), S. 176.

³²⁶ Ein derartiger Vergleich erfolgt, indem zunächst bedeutungslose Unterschiede zwischen den Konfigurationen (z. B. durch Rotation und Größenunterschiede) mittels sogenannter prokrustischer Verfahren ausgeglichen werden und anschließend die Übereinstimmung anhand eines Kongruenzkoeffizienten quantifiziert wird, vgl. dazu *Borg/Staufenbiel* (2007), S. 175–177, *Borg/Groenen* (2005), S. 429–441 und *Borg* (2010), S. 17–18.

³²⁷ Eine Partition ist definiert als „Familie disjunkter nichtleerer Teilmengen einer Menge, die die gesamte Menge überdecken“, vgl. *Schichl/Steinbauer* (2012), S. 150.

näherungsweise als Partition angesehen werden. Ein Vergleich könnte somit anhand der Partitionen erfolgen. Dazu können Übereinstimmungsmaße des Sorting-Verfahrens bzw. der (hierarchischen) Clusteranalyse entliehen werden, die dort verwendet werden, um die Übereinstimmung zwischen zwei Sortierungen bzw. zwischen zwei Lösungen der (hierarchischen) Clusteranalyse zu bestimmen.

5.3 Vergleich von Partitionen

Bei dem Vergleich von Partitionen wird überprüft, ob sich die $Z(Z-1)/2$ Objektpaare in gemeinsamen oder in unterschiedlichen Clustern befinden. Für den Vergleich von zwei Partitionen, die im Folgenden mit A und B bezeichnet werden, ergeben sich vier verschiedene Kombinationen, die in Tab. 5.1 dargestellt sind. Die Anzahl der Objektpaare in den jeweiligen Kombinationen wird mit a , b , c und d bezeichnet, wobei $a + b + c + d = Z(Z-1)/2$ gilt. Bei a und d handelt es sich um die Anzahl der „äquivalenten“ Paare und bei b und c um die Anzahl der „diskrepananten“ Paare.³²⁸

Tab. 5.1: Objektpaare zum Vergleich von Partitionen

| | | A | |
|-----|---|---------------------------------|---|
| | | Objektpaare im gleichen Cluster | Objektpaare in unterschiedlichen Clustern |
| B | Objektpaare im gleichen Cluster | a | b |
| | Objektpaare in unterschiedlichen Clustern | c | d |

Quelle: In Anlehnung an Milligan/Cooper (1986, S. 443).

Ausgehend von den vier Kombinationen lassen sich unterschiedliche Übereinstimmungsmaße bestimmen. Ein weit verbreitetes Übereinstimmungsmaß stellt der sogenannte Rand-Index dar, der auch als Simple Matching Koeffizient bezeichnet wird. Er beschreibt das Verhältnis

³²⁸ Vgl. Bortz (2005), S. 582.

zwischen der Anzahl äquivalenter Paare und der Anzahl aller Objektpaare:

$$RI = \frac{a+d}{a+b+c+d}. \quad (5.10)$$

Bei diesem Übereinstimmungsmaß erfolgt keine Unterscheidung zwischen übereinstimmender Gleichheit und übereinstimmender Ungleichheit.³²⁹ Der Term im Nenner ist konstant und ermöglicht eine Normierung des Rand-Indexes auf den Wertebereich zwischen 0 und 1. Ein Wert von 0 liegt vor, wenn keine Übereinstimmung zwischen zwei Partitionen besteht und zwar dergestalt, dass in der einen Partition alle Objekte ein gemeinsames Cluster bilden und in der anderen Partition jedes Objekt ein eigenes Cluster bildet. Ein Wert von 1 liegt bei zwei identischen Partitionen vor.³³⁰

Eine Übertragung dieses Übereinstimmungsmaßes, um jeweils eine Partition der Sorting-Daten auf Individualebene mit der Partition der interpretierten MDS-Konfiguration zu vergleichen, ist möglich.³³¹ Auf diese Weise lässt sich für jeden individuellen Datensatz jeweils ein Rand-Index bestimmen. Dabei bietet der Rand-Index im Gegensatz zu anderen Übereinstimmungsmaßen den Vorteil, dass aufgrund des konstanten Terms im Nenner das Übereinstimmungsmaß ausschließlich von der Anzahl der äquivalenten Paare abhängt. Insgesamt kann für alle Datensätze eine Verteilung der Rand-Indizes bestimmt werden. Die Auswahl zwischen mehreren konkurrierenden Interpretationen kann vor dem Hintergrund der Verteilung der Rand-Indizes erfolgen.

³²⁹ Vgl. zu Varianten des Rand-Indexes und anderen Übereinstimmungsmaßen *Coxon* (1999), S. 36–41, *Milligan/Cooper* (1986), S. 443–444, *Hubert/Arabie* (1985) und *Bortz* (2005), S. 581–583 sowie die dort zitierte Literatur.

³³⁰ Vgl. *Rand* (1971), S. 847.

³³¹ Bei dieser Übertragung bezieht sich der in diesem Kapitel verwendete Begriff „Cluster“ gleichermaßen auf gebildete Gruppen beim Sorting-Verfahren und auf Cluster bzw. klar voneinander trennbare Regionen einer Konfiguration der multidimensionalen Skalierung.

5.4 Anwendungsbeispiel zur empirischen Entwicklung von Konstrukten

Bei fast allen Anwendungen des Verfahrens der (integrierten) hierarchischen Informationsintegration wurde auf eine empirische Entwicklung von Konstrukten verzichtet. Eine Ausnahme bilden *Bos/Molin/van der Heijden* (2002), *Bos et al.* (2003) und *Bos* (2004) mit der empirischen Entwicklung von Konstrukten zur Beschreibung von P+R Anlagen in HII-Experimenten. Zur Erhebung von (Un-)Ähnlichkeitsdaten in Bezug auf 30 Objekte wurde ein Sorting-Verfahren eingesetzt. Dabei handelte es sich um ein Hybrid aus den Verfahren des Free-Sortings und des Fixed-Sortings dergestalt, dass die Anzahl der zu bildenden Gruppen zwar nicht vorgegeben, aber auf maximal sechs Gruppen begrenzt wurde. Diese Begrenzung der Gruppenanzahl wurde mit der übersichtlicheren Darstellung am Computerbildschirm technisch begründet. Zudem zeigte ein Pretest ohne Begrenzung der Gruppenanzahl, dass bei diesem Anwendungsfall Befragte selten mehr als sechs Gruppen bildeten.³³²

In einer internetbasierten Datenerhebung wurden (Un-)Ähnlichkeitsdaten von insgesamt 577 Befragten erhoben. Diese Unähnlichkeitsdaten wurden aggregiert und mit einem Verfahren der multidimensionalen Skalierung unter Verwendung des *Alscal*-Algorithmus ausgewertet. Dabei wurden Konfigurationen mit unterschiedlicher Dimensionalität ermittelt und schließlich eine Konfiguration mit drei Dimensionen ausgewählt. In dieser Konfiguration wurden fünf Cluster, die teilweise aus Subclustern bestehen, identifiziert.

Zudem wurden differenziert nach soziodemographischen Kriterien der Befragten Teilgruppen gebildet und für die jeweiligen aggregierten Unähnlichkeitsdaten Konfigurationen ermittelt. In einem optischen Vergleich wurden nur geringfügige Unterschiede zwischen den Konfigurationen der Teilgruppen festgestellt.³³³ Eine Überprüfung, ob die

³³² Da die Gruppenanzahl wesentlich von den zu sortierenden Objekten abhängt, ist eine verallgemeinernde Aussage bezüglich einer Begrenzung der Gruppenanzahl nicht möglich. Würde ein Befragter beim Sorting-Verfahren normalerweise mehr Gruppen verwenden, zwingt ihn eine solche Begrenzung zu einer anderen Gruppenbildung, was eine Veränderung der Ähnlichkeitsdaten zur Folge hat.

³³³ Vgl. *Bos* (2004), S. 53–54.

interpretierte Konfiguration der aggregierten Unähnlichkeitsdaten auch die Sorting-Daten auf Individualebene geeignet abbildet, erfolgte nicht.

Das Sorting-Verfahren mit einer entsprechenden Auswertung wurde erstmalig von *Bos/Molin/van der Heijden* (2002), *Bos et al.* (2003) und *Bos* (2004) zur Entwicklung von Konstrukten für Experimente der (integrierten) hierarchischen Informationsintegration eingesetzt. Grundsätzlich neu war dieses Verfahren jedoch nicht und wurde vorher bereits in unterschiedlichen Anwendungsbereichen eingesetzt.³³⁴ Dazu zählen beispielsweise die Kategorisierung von Charaktereigenschaften³³⁵, verbalen Konzepten³³⁶, sozialen Gruppen³³⁷, Verwandtschaftsbeziehungen³³⁸ und Nahrungsmitteln³³⁹.

Neben der Verwendung des Sorting-Verfahrens zu wissenschaftlichen Zwecken wird dieses Verfahren auch im sogenannten Usability Engineering eingesetzt und in diesem Zusammenhang als Card-Sorting bezeichnet. Dabei werden die Ergebnisse beispielsweise eingesetzt um die Struktur einer Webseite oder einer Software so zu gestalten, dass die Bedienung durch einen Nutzer möglichst intuitiv erfolgen kann.³⁴⁰

5.5 Zusammenfassung

In den Experimenten der (integrierten) hierarchischen Informationsintegration sollen durch die Konstrukte (ähnliche) Attribute zusammengefasst werden. Zur empirischen Entwicklung der Konstrukte bieten sich Verfahren zur Auswertung von (Un-)Ähnlichkeitsdaten an. Unterschiedliche Verfahren können zur Erhebung von (Un-)Ähnlichkeitsdaten eingesetzt werden. Im Rahmen dieser Arbeit wurde das Verfahren des Free-Sortings ausgewählt, bei dem Befragte die Objekte³⁴¹ dergestalt in nicht überlappende Gruppen sortieren, dass sich jeweils ähnliche Objekte in

³³⁴ Vgl. für eine umfangreiche Auflistung von Anwendungen des Sorting-Verfahrens *Coxon* (1999), S. 85–87.

³³⁵ Vgl. *Rosenberg/Nelson Carnot/Vivekananthan* (1968).

³³⁶ Vgl. *Miller* (1969).

³³⁷ Vgl. *Jones/Ashmore* (1973).

³³⁸ Vgl. *Rosenberg/Park Kim* (1975).

³³⁹ Vgl. *Lawless/Sheng/Knoops* (1995).

³⁴⁰ Vgl. *Spencer* (2009), S. 3–4 und *Kurniawan/Zaphiris* (2003).

³⁴¹ Zur Abgrenzung wird der Begriff Objekt anstelle von Attribut verwendet, vgl. Fußnote 275.

einer Gruppe befinden. Diese Daten lassen sich in einer Ähnlichkeitsmatrix (bzw. Unähnlichkeitsmatrix) ausdrücken, welche die Häufigkeit des gemeinsamen Auftretens (bzw. des gemeinsamen Nicht-Auftretens) angibt. Im Vergleich zu den anderen Verfahren eignet sich das Sorting-Verfahren insbesondere bei einer großen Anzahl von zu untersuchenden Objekten. Da jedoch nicht die Stärke der (Un-)Ähnlichkeit zwischen Objektpaaren ausgedrückt wird, ist – im Gegensatz zu anderen Verfahren – keine differenzierte Auswertung der Daten auf Individualebene möglich. Aufgrund der Entwicklung von einheitlichen Konstrukten ist im Rahmen der vorliegenden Arbeit eine Analyse auf Individualebene nicht erforderlich, sodass das Verfahren des Free-Sortings adäquat erscheint.

Zur Auswertung der Daten bietet sich ein Verfahren der multidimensionalen Skalierung (MDS) an. Dabei werden die Objekte in einem mehrdimensionalen Raum als Objektpunkte so dargestellt, dass die Distanzen zwischen den Objektpunkten bestmöglich den Unähnlichkeitsdaten entsprechen. Ähnliche Objekte liegen in dieser sogenannten Konfiguration näher beieinander als unähnliche Objekte. Gemäß dem Skalenniveau der Daten der aggregierten Unähnlichkeitsmatrix kann eine metrische MDS für ratioskalierte Daten eingesetzt werden. Mit unterschiedlichen Maßen der Anpassungsgüte kann beschrieben werden, wie gut die Distanzen die aggregierten Daten abbilden. Den Wert eines solchen Anpassungsmaßes gilt es bei der Bestimmung einer geeigneten Konfiguration iterativ zu minimieren, wobei unterschiedliche Algorithmen verwendet werden können.

Die Koordinatenachsen der unterschiedlichen Dimensionen der Konfiguration haben in der Regel keine inhaltlich interpretierbare Bedeutung, wesentlich ist hingegen die relative Lage der Objektpunkte zueinander. Für die Entwicklung von Konstrukten ist die Identifikation sogenannter Cluster von Interesse, d. h. Gruppen mit nah beieinander liegenden Objekten, wobei zwischen solchen Gruppen leerer Raum liegt. Eine derartig interpretierte Konfiguration kann näherungsweise als Partition im mathematischen Sinne angesehen werden.

Da die Konfiguration auf aggregierten Daten basiert, werden möglicherweise die Unähnlichkeitsdaten auf Individualebene nicht geeignet abgebildet. Eine Überprüfung ist durch einen Vergleich der Partition der inter-

pretierten Konfiguration mit den jeweiligen Partitionen der Sorting-Daten auf Individualebene möglich. Dieser Vergleich erfolgt anhand von Objektpaaren. Ein mögliches Übereinstimmungsmaß stellt der Rand-Index dar, welcher das Verhältnis zwischen der Anzahl äquivalenter Paare (übereinstimmende Gleichheit bzw. Ungleichheit) und der Anzahl aller Objektpaare beschreibt. Somit lässt sich ein Rand-Index für jeden einzelnen Datensatz auf Individualebene ermitteln. Eine Entscheidung zwischen konkurrierenden Interpretationen kann vor dem Hintergrund der Verteilung der Rand-Indizes getroffen werden.

Zwar ist eine (solche) empirische Entwicklung von Konstrukten für Experimente der (integrierten) hierarchischen Informationsintegration bislang selten. In anderen Anwendungsbereichen wurde das Sorting-Verfahren mit entsprechender Auswertung jedoch schon vielfach eingesetzt.

6 Empirische Entwicklung von Konstrukten für die Anwendung des Verfahrens der HII-I im Regionalverkehr

Ein Ziel dieser Arbeit ist die Analyse des Einflusses der Qualität im regionalen ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr. Da eine Analyse dieses Zusammenhangs anhand von Realdaten nur sehr eingeschränkt möglich ist, ist eine Befragung zur Erhebung von Experimentaldaten erforderlich. Grundsätzlich können zu diesem Zweck gängige Wahlexperimente eingesetzt werden. Jedoch ist dabei die Anzahl der Attribute zur Charakterisierung der Alternativen aufgrund der Komplexität begrenzt. Wahlexperimente unter Verwendung des Verfahrens der integrierten hierarchischen Informationsintegration (vgl. Kap. 4.) erscheinen daher geeigneter.

In Befragungen insbesondere von Nutzern, aber auch von Nicht-Nutzern des regionalen SPNV sollen Experimente der integrierten hierarchischen Informationsintegration durchgeführt werden, in denen als Wahlalternativen eine Regionalbahn, ein Regionalbus und ggf. ein Pkw zur Verfügung stehen. Für die Gestaltung der Wahlexperimente ist die Entwicklung von Konstrukten zur Beschreibung der öffentlichen Verkehrsmittel erforderlich. Die Entwicklung der Konstrukte erfolgt im Rahmen einer empirischen Vorstudie, die in diesem Kapitel beschrieben wird.³⁴² Die Anwendung der HII-I in den Verkehrsmittelwahlexperimenten erfolgt im Rahmen einer Hauptstudie, die in Kap. 7 beschrieben wird.

Da in der Hauptstudie zur Befragung der Nutzer und Nicht-Nutzer des regionalen SPNV möglichst einheitliche Experimente verwendet werden sollen, ist es erforderlich, dass die Konstrukte gleichermaßen für Nutzer als auch Nicht-Nutzer des regionalen SPNV geeignet sind. Die Entwicklung der Konstrukte erfolgt unter Verwendung des Verfahrens des Free-Sortings zur Erhebung von (Un-)Ähnlichkeitsdaten, welche in aggregier-

³⁴² Vgl. für die Beschreibung der Vorstudie auch *Richter/Keuchel* (2012a).

ter Form mit einem Verfahren der multidimensionalen Skalierung für ratioskalierte Daten ausgewertet werden.

Allgemein sind bei dem Verfahren des Free-Sortings Objekte in Bezug auf ihre Ähnlichkeit zu beurteilen. Im vorliegenden Anwendungsfall handelt es sich bei den in Rede stehenden Objekten um Qualitätsmerkmale von Regionalbahnen und Regionalbussen, die potentiell als (Qualitäts-)Attribute zur Charakterisierung der beiden ÖV-Alternativen in den Wahlexperimenten verwendet werden können.

Derartige Qualitätsmerkmale galt es zunächst zu identifizieren.³⁴³ Zu diesem Zweck wurde eine Literaturrecherche zu bereits existierenden Kundenzufriedenheits- und Qualitätsmessinstrumenten durchgeführt, die durch Gespräche mit Mitarbeitern von drei nichtbundeseigenen Eisenbahnverkehrsunternehmen³⁴⁴ ergänzt wurde. Dabei handelt es sich um die Keolis Deutschland GmbH & Co. KG Niederlassung eurobahn, die NordWestBahn GmbH und die WestfalenBahn GmbH. Diese Unternehmen stellen Kooperationspartner des Forschungsprojekts „Qualitätssimulationsmodell im Regionalverkehr“ dar, welches an der Westfälischen Hochschule (ehemals Fachhochschule Gelsenkirchen) durchgeführt und vom Bundesministerium für Bildung und Forschung im Rahmen der Förderlinie „FHprofUnd“ gefördert wurde.

Im Ergebnis wurde eine Liste mit 39 Qualitätsmerkmalen, welche potentiell einen Einfluss auf die Verkehrsmittelwahl ausüben, erstellt. Nach einem ersten Pretest, bei dem insbesondere das Erhebungsinstrument überprüft wurde, wurde diese Liste leicht modifiziert und umfasste schließlich 32 Qualitätsmerkmale, die in Tab. 6.1 dargestellt sind.

³⁴³ Vgl. *Bühner* (2006), S. 47–48 und *Gläser/Laudel* (2009), S. 74–76.

³⁴⁴ Empirische Untersuchungen deuten auf Unterschiede bezüglich der Beurteilung von Qualitätsmerkmalen von Kunden einerseits und Unternehmensvertretern andererseits hin, vgl. *Schnippe* (1999), S. 53–55. Die Unterschiede beziehen sich auf die nach Wichtigkeit geordnete Rangfolge. Im Rahmen der empirischen Entwicklung von Konstrukten erfolgte jedoch nur eine grobe Vorauswahl von Qualitätsmerkmalen.

Tab. 6.1: Qualitätsmerkmale zur empirischen Entwicklung von Konstrukten

| Qualitätsmerkmale | |
|--|--|
| <ul style="list-style-type: none"> • Kein Fahrtenausfall • Pünktlichkeit • Gesamtreisezeit (von Tür zu Tür) • Frequenz (Zahl der angebotenen Fahrten pro Stunde) • Umsteigehäufigkeit • Anschlussgarantie • Fahrpreis • Sitzplatzverfügbarkeit im Zug • Komfort der Sitzplätze • Sauberkeit der Toilette im Zug • Abstellmöglichkeiten im Zug (z. B. für Gepäck, Fahrräder, Kinderwagen) • Unterstellplatz/Wartehaus am Bahnsteig • Sauberkeit Zug innen • Sauberkeit Zug außen • Fahrplaninformationen am Bahnsteig • Fahrplanauskunft im Internet • Durchsagen im Zug bei Störungen oder Verspätungen | <ul style="list-style-type: none"> • Durchsagen am Bahnsteig bei Störungen oder Verspätungen • Informationsanzeige am Bahnsteig bei Störungen oder Verspätungen • Informationsanzeige am Zug (außen) • Informationsanzeige im Zug zu Anschlusszügen • Durchsage im Zug zu Anschlusszügen • Videoüberwachung im Zug • Videoüberwachung am Bahnsteig • Sicherheitspersonal im Zug • Sicherheitspersonal am Bahnsteig • Freundlichkeit des Personals • Fahrkartenautomat am Bahnsteig • Fahrkartenautomat im Zug • Fahrkartenverkauf durch Personal am Bahnhof • Fahrkartenverkauf durch Personal im Zug • Verständlichkeit des Tarifsystems |

Quelle: Eigene Darstellung.

6.1 Untersuchungsdesign und Datenerhebung

Die empirische Untersuchung wurde als computergestützte Face-to-face-Befragung³⁴⁵ konzipiert. Zwar ist allgemein dieses Erhebungsverfahren gegenüber schriftlichen oder internetgestützten Befragungen vergleichsweise kostenintensiv, doch bietet es den Vorteil, dass durch den Einsatz

³⁴⁵ Computergestützte Face-to-face Befragungen werden auch als „Computer Assisted Personal Interview“ (CAPI) bezeichnet, vgl. Schnell/Hill/Esser (2011), S. 368 und Weis/Steinmetz (2005), S. 116.

eines Interviewers auch längere Befragungen und durch den Einsatz von Laptops differenziertere Fragestellungen möglich sind.³⁴⁶

Das für die Datenerhebung erforderliche Erhebungsinstrument wurde mit *Visual Basic for Applications (VBA)* in *MS Access 2007* programmiert. Dem Erhebungsinstrument lag ein Fragebogen zugrunde, der aus fünf Teilen bestand:

1. Fragen zum allgemeinen Verkehrsverhalten,
2. Sortierungsaufgabe unter Verwendung des Verfahrens des Disjoint-Free-Sortings (vgl. Kap. 5.1),
3. Bewertungsaufgabe,
4. Soziodemografische Fragen und Danksagung und
5. Anmerkungen des Interviewers.

Eine Übersicht der einzelnen Fragebogenmasken befindet sich in Anhang B.

Die Fragen zum allgemeinen Verkehrsverhalten (Abb. C.1) bezogen sich auf den Fahrtzweck von hauptsächlich durchgeführten Fahrten, den Führerscheinbesitz, die Pkw-Verfügbarkeit, den Besitz einer Monatskarte³⁴⁷, die Nutzungshäufigkeit der Bahn sowie die Nutzungshäufigkeit des Pkw (auch als Mitfahrer). Alle Fragen wurden als „geschlossene“ Fragen formuliert.

Bei der Sortierungsaufgabe (Abb. C.2) waren die Qualitätsmerkmale so im Hinblick auf die wahrgenommene inhaltliche Ähnlichkeit in Gruppen zu sortieren, dass sich die Merkmale innerhalb einer Gruppe ähnlich waren und sich von den Merkmalen anderer Gruppen unterschieden. Die 32 zu sortierenden Qualitätsmerkmale waren von Beginn an sichtbar und in zufälliger Reihenfolge auf der linken Bildschirmseite angeordnet.³⁴⁸ Die rechte Bildschirmseite war zu Beginn der Sortierungsaufgabe leer, da weder die Anzahl noch die Bezeichnung der Gruppen vorgegeben war. Die einzelnen Merkmale ließen sich mittels „Drag & Drop“ auswählen

³⁴⁶ Vgl. Scholl (2009), S. 59–60, Scheffler (2000), S. 70–72, Schnell/Hill/Esser (2011), S. 368–369, Bortz/Döring (2006), S. 236–261, Gläser/Laudel (2009), S. 153–154, Jankisz/Moosbrugger (2007), S. 35–36 und Weis/Steinmetz (2005), S. 124.

³⁴⁷ Als Monatskarte wurden allgemein alle Zeitfahrtscheine sowie vergleichbare Fahrtberechtigungen zusammengefasst, darunter auch Semestertickets und Schwerbehindertenausweise.

³⁴⁸ Durch die zufällige Anordnung der Merkmale sollten systematische Reihenfolge- und Ermüdungseffekte vermieden werden.

und auf die rechte Bildschirmseite ziehen. Dadurch wurde entweder eine neue Gruppe erstellt oder ein Merkmal einer bereits bestehenden Gruppe zugeordnet. Die Sortierungsaufgabe war abgeschlossen, wenn alle Merkmale eindeutig zugeordnet waren. Zudem war für jede Gruppe eine frei wählbare Bezeichnung anzugeben; Zweck dieser Teilaufgabe war es, dass die Befragten die Gruppenbildung reflektierten. Änderungen der Gruppenbildung waren jederzeit – auch noch während der Bezeichnung – möglich. Für die Umsetzung der Sortierungsaufgabe wurde die internetbasierte³⁴⁹ Anwendung *WebSort*³⁵⁰ verwendet, die aus *MS Access* heraus gestartet wurde.

Bei der Bewertungsaufgabe (Abb. C.3 bis Abb. C.6) waren alle 32 Qualitätsmerkmale auf einer Rating-Skala dahingehend zu bewerten, wie wichtig das jeweilige Merkmal allgemein in Bezug auf die eigene Verkehrsmittelwahl war. Die Rating-Skala bestand aus sieben numerisch bezeichneten Stufen, wobei die erste Stufe mit *sehr unwichtig* und die höchste Stufe mit *sehr wichtig* verankert wurden. Darüber hinaus konnten in einer offenen Frage weitere Qualitätsmerkmale ergänzt und ebenfalls bewertet werden. Diese offene Frage diente als Kontrolle, dass keine wesentlichen Merkmale in der Vorstudie ausgelassen wurden.³⁵¹ Die Bewertungsaufgabe wurde bewusst erst im Anschluss an die Sortierungsaufgabe durchgeführt, um nicht durch die Reihenfolge der zu bewertenden Merkmale die Sortierung zu beeinflussen.³⁵²

Die soziodemografischen Fragen (Abb. C.7) bezogen sich auf Alter und Geschlecht des Befragten sowie die Anzahl der Personen insgesamt und der davon minderjährigen Personen in seinem Haushalt.

Die Interviewer hatten die Möglichkeit, im Anschluss an jedes Interview Anmerkungen (Abb. C.8) zu hinterlassen.

³⁴⁹ Die Herstellung der Internetverbindung erfolgte über das Mobilfunknetz. Bei einer temporären Verbindungsunterbrechung konnten die Daten nachträglich eingepflegt werden.

³⁵⁰ *WebSort* (www.websort.net) ist eine Anwendung zum „Online Card-Sorting“ des Unternehmens Lime & Chile Productions (Umbenennung des Unternehmens in UX Punk im Jahr 2011).

³⁵¹ Vgl. Fußnote 344.

³⁵² Eine andere Möglichkeit wäre gewesen, die Reihenfolge der zu bewertenden Merkmale zufällig zu variieren.

Zu einer groben Überprüfung des Erhebungsinstruments wurde ein Pretest³⁵³ in Form einer Befragung von insgesamt 24 Mitarbeitern und Studierenden der Fachhochschule Gelsenkirchen im Zeitraum 27.11.2009 – 5.12.2009 durchgeführt. Aufgrund einer relativ langen Interviewdauer (durchschnittliche Dauer³⁵⁴ der Interviews: 19:29 min, Standardabweichung (SD): 7:39 min) wurde die Anzahl der Qualitätsmerkmale mithilfe von Expertengesprächen und vor dem Hintergrund der Relevanz für diese Anwendung von 39 auf 32 reduziert. Darüber hinaus wurden geringfügige Anpassungen der Bezeichnungen vorgenommen. Die Daten dieses Pretests wurden nicht in die spätere Datenauswertung einbezogen.

Ein weiterer Pretest wurde in Zügen der Linie RE 3 (Dortmund – Düsseldorf) ab dem 26.01.2010 durchgeführt. Der Übergang zwischen diesem zweiten Pretest und der eigentlichen Datenerhebung war fließend, da keine weiteren Anpassungen des Erhebungsinstruments vorgenommen wurden. Diese Daten wurden somit in die spätere Datenauswertung einbezogen.

Im Rahmen der Vorstudie erfolgte die Befragung der Nutzer des regionalen SPNV in den Zügen der Kooperationspartner. Fahrgastbefragungen in den Verkehrsmitteln bieten den Vorteil, dass Fahrgäste in der Regel Wartezeiten überbrücken, sodass eine hohe Bereitschaft zur Befragungsteilnahme erwartet wurde.

Idealerweise wird eine derartige Befragung unter Verwendung eines geschichteten Stichprobenverfahrens (probabilistisches Stichprobenverfahren) durchgeführt. Dabei erfolgt in der ersten Stufe die Auswahl von „Wagenfahrten“. In der zweiten Stufe werden Fahrgäste zufällig ausgewählt, die sich in den diesen „Wagenfahrten“ entsprechenden Zügen befinden. Mit den Daten einer (parallel stattfindenden) Fahrgastzählung werden die erhobenen Daten gewichtet.³⁵⁵ Unabhängig vom verwendeten

³⁵³ Vgl. Schnell/Hill/Esser (2011), S. 340–344, Gläser/Laudel (2009), S. 150 und Bortz/Döring (2006), S. 248.

³⁵⁴ Da die Zeit erst nach dem Ausfüllen des ersten Fragebogenteils protokolliert wurde, beinhaltet diese Interviewdauer nicht die Zeit zur Beantwortung der Fragen zum allgemeinen Verkehrsverhalten.

³⁵⁵ Vgl. Pfeiffer/Schmidt (2006), S. 66, Bortz/Döring (2006), S. 402 und Schnell/Hill/Esser (2011), S. 271–272.

Stichprobenverfahren besteht jedoch die Problematik, dass manche Fahrgäste die Teilnahme an einer Befragung verweigern.³⁵⁶

Bedingt dadurch, dass detaillierte Fahrgastzählungen vertrauliche Daten der Verkehrsunternehmen darstellen, im Rahmen dieser Arbeit ein Schwerpunkt auf der Überprüfung der hierarchischen Struktur der HII-I liegt und darüber hinaus forschungsökonomische Gründe eine Rolle spielten, erfolgte keine Fahrgastzählung und die Datenerhebung wurde in Abhängigkeit der zeitlichen Verfügbarkeit³⁵⁷ der Interviewer durchgeführt. Es handelt sich somit um eine Ad-hoc-Stichprobe (nichtprobabilistisches Verfahren).³⁵⁸ Die Aussagekraft einer solchen Stichprobe ist grundsätzlich eingeschränkt. Jedoch wurden bei der Datenerhebung der genaue Erhebungszeitpunkt und der Erhebungsort (Liniennummer und Fahrtrichtung) dokumentiert, sodass ein Rückschluss auf „Wagenfahrten“ möglich wäre. Unter Kenntnis der (durchschnittlichen) Fahrgastzahlen der entsprechenden „Wagenfahrten“ wäre eine Gewichtung möglich.

Die Befragung der Nutzer des regionalen SPNV erfolgte auf den Linien RE 3 (Dortmund – Düsseldorf), RB 61 (Bielefeld – Bad Bentheim), RB 66 (Münster – Osnabrück) und RB 67 (Bielefeld – Münster) im Zeitraum 26.01. – 15.04.2010.

Da die Nicht-Nutzer des SPNV definitionsgemäß nicht in Zügen anzutreffen sind, musste zur Befragung dieser Zielgruppe ein anderer Erhebungsort ausgewählt werden. Da die Durchführung von Hausinterviews (z. B. in der Privatwohnung oder an der Arbeitsstätte) im Allgemeinen sehr kosten- und zeitintensiv ist, wurde stattdessen eine Befragung im öffentlichen Raum (Passanteninterviews) angestrebt.³⁵⁹ Eine vergleichsweise erfolgreiche Akquisemöglichkeit wurde in Wartebereichen von Behörden (z. B. Kfz-Zulassungsstelle und Bürgerämter) gesehen. Im Zeitraum 03.05. – 06.05.2010 erfolgte eine Befragung von Wartenden in den Kfz-Zulassungsstellen in Bielefeld und Osnabrück. Im

³⁵⁶ Vgl. *Dillman et al.* (2002), S. 6–10, *Schnell* (1997), S. 119–122 und *Schnell/Hill/Esser* (2011), S. 305–307.

³⁵⁷ Dabei wurde dennoch darauf geachtet, dass an Werktagen (exkl. Samstag) tagsüber zu vielen unterschiedlichen Zeiten Daten erhoben wurde.

³⁵⁸ Vgl. *Bortz/Döring* (2006), S. 402 und *Schnell/Hill/Esser* (2011), S. 291.

³⁵⁹ Vgl. *Scholl* (2009), S. 29–30.

Rahmen der Vorstudie diente diese Gruppe von Befragten als Kontrollgruppe.

Zur Unterstützung bei der Datenerhebung wurden studentische Interviewer eingesetzt. Diese wurden auf ihren Einsatz in einer Schulung³⁶⁰ vorbereitet, in der über den Gegenstand der Befragung, Aufbau des Fragebogens und Bedienung des Erhebungsinstruments ausführlich informiert wurde. Die Interviewer wurden insbesondere angeleitet, explizit darauf hinzuweisen, dass die Sortierung der Qualitätsmerkmale im Hinblick auf die wahrgenommene inhaltliche Ähnlichkeit erfolgen sollte (und nicht im Hinblick auf beispielsweise Wichtigkeit oder Vorhandensein). Die Interviewsituation wurde in Probeinterviews geübt.

Insgesamt wurden 509 gültige Interviews³⁶¹ durchgeführt, davon 477 in Zügen des Regionalverkehrs und 32 in Kfz-Zulassungsstellen. Die Interviews dauerten durchschnittlich 11:46 min (SD: 4:29 min).³⁶²

6.2 Beschreibung der Stichprobe

Die soziodemografischen Fragen und Fragen zum allgemeinen Verkehrsverhalten wurden getrennt nach den beiden Erhebungsorten ausgewertet. Eine Stichprobenbeschreibung mit absoluten und adjustierten relativen Häufigkeiten³⁶³ ist in Tab. 6.2 dargestellt.

Die Befragten in den Zügen waren zwischen 15 und 76 Jahren alt, wobei die Alterskategorie *18 – 30 Jahre* mit 52,7 % die größte Gruppe bildete. Das Durchschnittsalter betrug 33,3 Jahre (SD: 12,8 Jahre, Median: 29 Jahre). Mehr als die Hälfte der Befragten lebte in Ein- oder Zweipersonenhaushalten und nur in jedem vierten Haushalt lebte eine minderjährige Person. Mit 54,9 % war die Mehrheit der Befragten in den

³⁶⁰ Vgl. Bortz/Döring (2006), S. 247–248 und Schnell/Hill/Esler (2011), S. 344–345.

³⁶¹ Interviews werden hierbei als gültig bezeichnet, wenn die Sortierungsaufgabe vollständig durchgeführt wurde. Ausgenommen sind jedoch ca. 5 % der insgesamt erhobenen Datensätze, bei denen offensichtlich die Befragten trotz entsprechender Hinweise die Aufgabestellung (Sortierung nach inhaltlicher Ähnlichkeit) nicht erfüllt haben. Diese Datensätze wurden anhand der Anmerkungen der Interviewer bzw. der Gruppenbenennungen herausgefiltert.

³⁶² Auch diese Zeit beinhaltet nicht die Dauer der Beantwortung der Fragen zum allgemeinen Verkehrsverhalten, vgl. Fußnote 354.

³⁶³ Bei adjustierten relativen Häufigkeiten werden fehlende Angaben nicht berücksichtigt.

Tab. 6.2: Absolute und relative adjustierte Häufigkeiten der Charakteristika zur Soziodemografie und zum allgemeinen Verkehrsverhalten differenziert nach Nutzergruppen

| | Bahnnutzer (N=477) | | Pkw-Nutzer (N=32) | |
|--|-----------------------|---------------------|----------------------|---------------------|
| | Abs. | Rel. adj. (in %) | Abs. | Rel. adj. (in %) |
| Geschlecht | | | | |
| Männlich | 260 | 54,9 | 19 | 59,4 |
| Weiblich | 214 | 45,1 | 13 | 40,6 |
| Alter (Jahren) | | | | |
| ≤ 17 | 8 | 1,7 | 0 | 0,0 |
| 18 - 30 | 245 | 52,7 | 8 | 27,6 |
| 31 - 50 | 154 | 33,1 | 19 | 65,5 |
| ≥ 51 | 59 | 12,5 | 2 | 6,9 |
| Pkw-Führerscheinbesitz | | | | |
| Ja | 398 | 83,8 | 32 | 100,0 |
| Nein | 77 | 16,6 | 0 | 0 |
| Allg. Pkw-Verfügbarkeit | | | | |
| Immer | 177 | 37,1 | 28 | 87,5 |
| Meistens | 60 | 12,6 | 3 | 9,4 |
| Manchmal | 69 | 14,5 | 1 | 3,1 |
| Selten | 53 | 11,1 | 0 | 0,0 |
| Nie | 118 | 24,7 | 0 | 0,0 |
| Monatsticketbesitz | | | | |
| Ja | 349 | 73,5 | 6 | 18,8 |
| Nein | 126 | 26,5 | 26 | 81,3 |
| Allg. Nutzungshäufigkeit Bahn | | | | |
| Mehrmals in der Woche | 325 | 68,1 | 4 | 12,5 |
| Mehrmals im Monat | 88 | 18,4 | 6 | 18,8 |
| Selten | 63 | 13,2 | 16 | 50,0 |
| Nie | 1 | 0,2 | 6 | 18,8 |
| Allg. Nutzungshäufigkeit Pkw | | | | |
| Mehrmals in der Woche | 189 | 39,6 | 28 | 87,5 |
| Mehrmals im Monat | 114 | 23,9 | 4 | 12,5 |
| Selten | 126 | 26,4 | 0 | 0,0 |
| Nie | 48 | 10,1 | 0 | 0,0 |
| Anzahl der Personen im Haushalt | | | | |
| 1 | 121 | 26,2 | 4 | 13,8 |
| 2 | 138 | 29,9 | 7 | 24,1 |
| 3 | 83 | 18,0 | 9 | 31,0 |
| 4 | 74 | 16,1 | 6 | 20,7 |
| 5 und mehr | 45 | 9,8 | 3 | 10,3 |

Quelle: Eigene Erhebung.

Zügen männlich. Fast die Hälfte der Befragten gab den Beruf und mehr als ein Viertel die Ausbildung als Zweck für hauptsächlich durchgeführte Fahrten an. Mit 83,8 % verfügte ein wesentlicher Teil der Befragten in den Zügen über einen Pkw-Führerschein. Von den Befragten in den Zügen insgesamt hatten 37,1 % einen Pkw *immer* zur Verfügung, bezogen auf die Teilgruppe mit Führerschein 44,2 %. Fast dreiviertel der Befragten verfügten über ein Monatsticket³⁶⁴. Insgesamt nutzten 68,1 % den Zug *mehrmals pro Woche*, aber nur 39,6 % den Pkw (auch als Mitfahrer).

Die Befragten in den Kfz-Zulassungsstellen waren zwischen 18 und 69 Jahre alt. Mit 65,5 % gehörten die meisten Befragten der Alterskategorie *31 – 50 Jahre* an. Das Durchschnittsalter lag bei 36,5 Jahren (SD: 11,9 Jahre, Median: 37 Jahre). Von den Befragten lebten 37,9 % in Ein- oder Zweipersonenhaushalten und nur in jedem vierten Haushalt lebte eine minderjährige Person. Mit 59,4 % war die Mehrheit der Befragten männlich. Als Zweck für hauptsächlich durchgeführte Fahrten gaben 62,5 % der Befragten den Beruf an und nur 3,1 % die Ausbildung. Alle Befragten verfügten über einen Pkw-Führerschein. Mit 87,5 % hatte ein sehr großer Anteil der Befragten *immer* einen Pkw zur Verfügung. Ein Monatsticket besaßen 18,8 % der Befragten. Den Pkw nutzten 87,5 % der Befragten *mehrmals in der Woche*, jedoch nur 12,5 % die Bahn. Die Bahn wurde von 50 % der Befragten *selten* und von 18,8 % der Befragten sogar *nie* benutzt.

Aufgrund des Nutzungsverhaltens werden im Folgenden die Befragten in den Zügen auch vereinfachend als *Bahnnutzer* und die Befragten in den Kfz-Zulassungsstellen als *Pkw-Nutzer* bezeichnet.

6.3 Analyse der Unähnlichkeitsdaten

Zur Auswertung der Unähnlichkeitsdaten war zunächst eine Datenaufbereitung erforderlich. Die mit *WebSort* erhobenen Sorting-Daten, die online als Ähnlichkeitsmatrizen auf Individualebene gespeichert waren, wurden heruntergeladen, lokal gespeichert und mit den in *MS Access* erhobenen Daten verknüpft. Diese Verknüpfung ermöglichte eine

³⁶⁴ Der Begriff „Monatsticket“ wurde weit gefasst, vgl. Fußnote 347.

Differenzierung der Ähnlichkeitsdaten nach den erhobenen Merkmalen, u. a. dem Erhebungsort.

Die Analyse der Unähnlichkeitsdaten mit einem Verfahren der multi-dimensionalen Skalierung basiert ausschließlich auf den Daten der *Bahnnutzer*. Zu diesem Zweck wurden zunächst die entsprechenden Ähnlichkeitsmatrizen auf Individualebene in Unähnlichkeitsmatrizen auf Individualebene transformiert. Anschließend wurden diese 477 Unähnlichkeitsmatrizen durch Summation aggregiert. Die aggregierte Unähnlichkeitsmatrix ist in Anhang B (Tab. B.1) dargestellt.

Die Auswertung der Daten erfolgte in aggregierter Form in *SPSS* (Version 19 und 20)³⁶⁵ mit dem *Proxscal*-Algorithmus für ratioskalierte Daten. Als Abbruchkriterium wurden eine Stress-Konvergenz von 0,0001 und ein Mindeststress von ebenfalls 0,0001 sowie eine maximale Anzahl von 1000 Iterationen gewählt. Darüber hinaus wurden die Standardeinstellungen des Algorithmus beibehalten.³⁶⁶

Es wurden Konfigurationen mit eins bis acht Dimensionen ermittelt. Für jede Dimensionalität wurden drei unterschiedliche Ausgangskonfigurationen bestimmt (unter Verwendung des Simplex- und des Torgerson-Algorithmus sowie mehrfache Zufallsstarts (jeweils 100.000 Ziehungen)) und jeweils die Endkonfiguration mit dem niedrigsten normalisierten Rohstress ausgewählt.³⁶⁷ Für diese acht Endkonfigurationen sind jeweils die Werte des normalisierten Rohstresses und Stress-1 sowie deren Veränderungen in Tab. 6.3 dargestellt.

Wie in Kap. 5.2 beschrieben ist in manchen Fällen die Auswahl der Dimensionalität anhand eines Scree-Plots möglich. Für die ermittelten Konfigurationen wurde ein Scree-Plot mit den Werten des normalisierten Rohstresses erstellt, welcher in Abb. 6.1 dargestellt ist. Ein „Ellbogen“ zur Auswahl der Dimensionalität konnte nicht identifiziert werden.

³⁶⁵ Vgl. *IBM Corporation* (2010 und 2011).

³⁶⁶ Vgl. *IBM Corporation* (2011b), S. 1669–1670.

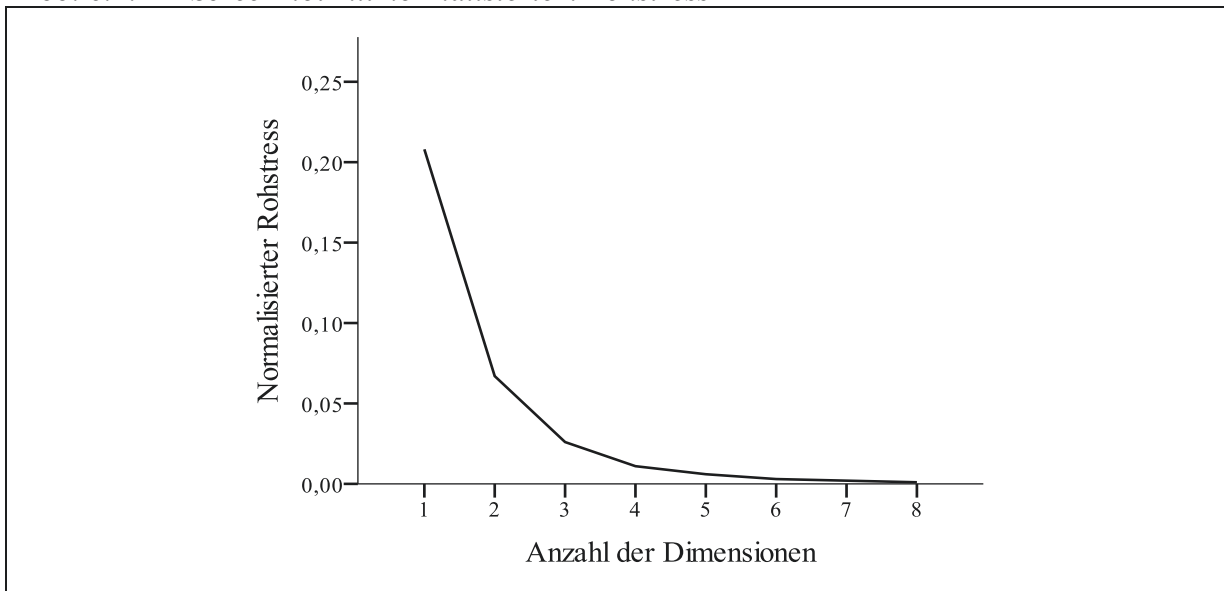
³⁶⁷ Bei der Konfiguration mit einer Dimension führte der Simplex-Algorithmus, bei zwei bis sieben Dimensionen die mehrfachen Zufallsstarts und bei acht Dimensionen der Torgerson-Algorithmus zu dem jeweils besten Gesamtergebnis. Dabei unterschieden sich die Werte des normalisierten Rohstress nur marginal, was auf die (geringen) Schwellenwerte der Abbruchkriterien im Rahmen des iterativen Optimierungsprozesses zurückzuführen ist.

Tab. 6.3: Anpassungsmaße der ermittelten Konfigurationen

| Anzahl der Dimensionen | Normalisierter Rohstress | Stress-1 | Veränderung des normalisierten Rohstresses | Veränderung des Stress-1 |
|------------------------|--------------------------|----------|--|--------------------------|
| 1 | 0,208 | 0,456 | 0,792 | 0,544 |
| 2 | 0,067 | 0,260 | 0,140 | 0,196 |
| 3 | 0,026 | 0,162 | 0,041 | 0,098 |
| 4 | 0,011 | 0,106 | 0,015 | 0,056 |
| 5 | 0,006 | 0,076 | 0,005 | 0,030 |
| 6 | 0,003 | 0,058 | 0,002 | 0,018 |
| 7 | 0,002 | 0,046 | 0,001 | 0,012 |
| 8 | 0,001 | 0,038 | 0,001 | 0,008 |

Quelle: Eigene Darstellung.

Abb. 6.1: Scree-Plot mit normalisiertem Rohstress



Quelle: Eigene Darstellung.

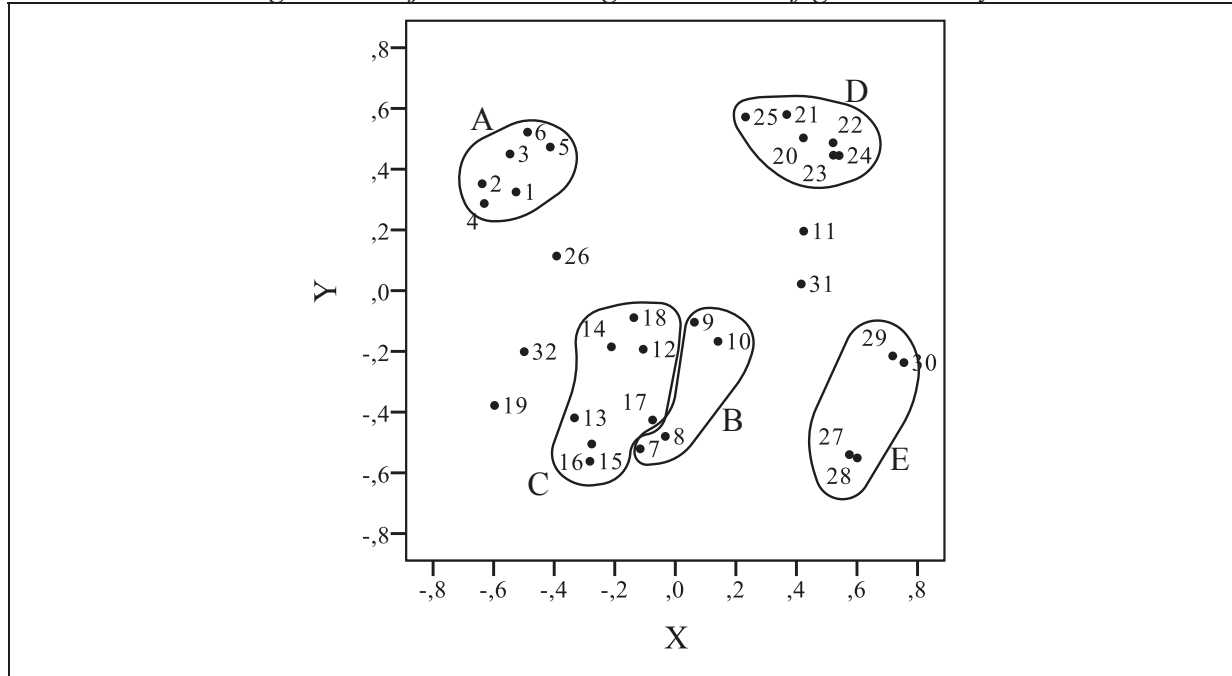
Daher wurde unter folgenden Aspekten die Konfiguration mit drei Dimensionen ausgewählt: Im Vergleich zur Konfiguration mit zwei Dimensionen verbessert sich der normalisierte Rohstress durch eine dritte Dimension noch substantiell. Eine vierte Dimension bewirkt nur eine vergleichsweise geringe Verbesserung und führt zudem zu Problemen bei der Interpretation der Konfiguration.

6.4 Interpretation und Überprüfung der Konfiguration

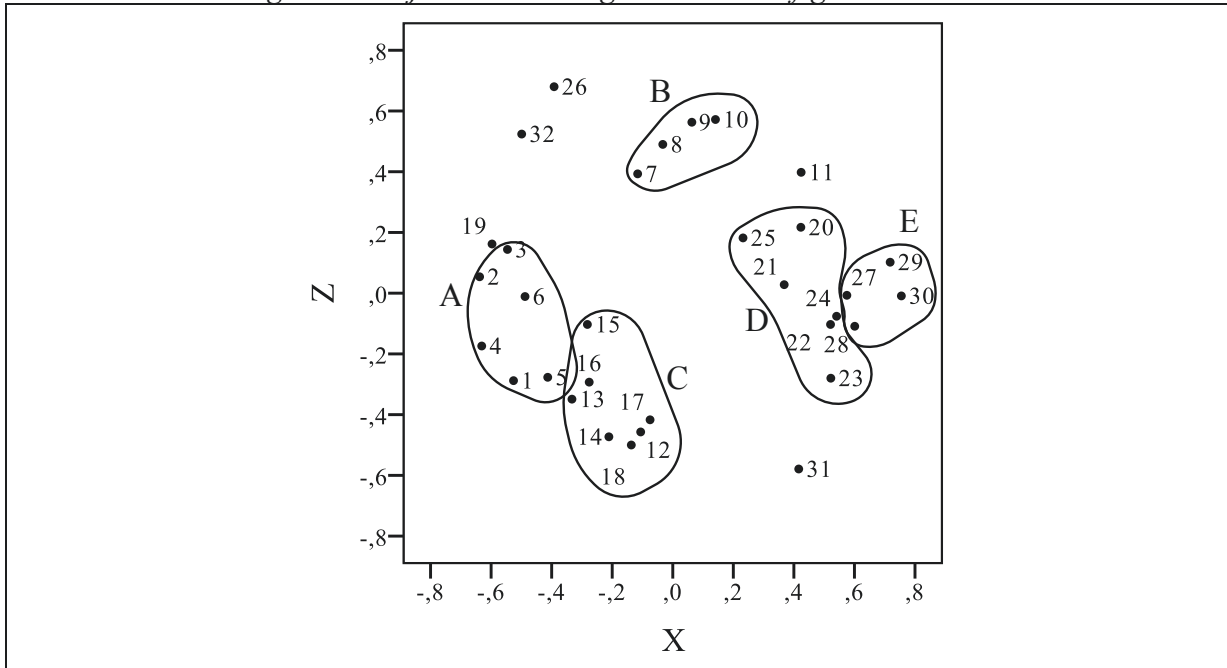
Bei einer Konfiguration mit drei Dimensionen lässt sich der aufgespannte Raum als Würfel vorstellen, innerhalb dessen sich die Objektpunkte, d. h. die 32 Qualitätsmerkmale befinden. Zur besseren Interpretierbarkeit wurden orthogonale Projektionen dieser Konfiguration erstellt, welche in Abb. 6.2 bis Abb. 6.4 dargestellt sind. Sie entsprechen dem Blick in einem rechten Winkel auf drei nicht gegenüberliegende Flächen des Würfels.

Die Koordinatenachsen dieser Konfiguration ließen sich, auch nach einer Rotation, nicht inhaltlich interpretieren. Allerdings wurden fünf Cluster identifiziert, welche in den orthogonalen Projektionen Abb. 6.2 bis Abb. 6.4 eingezeichnet und mit *Verbindungsqualität* (A), *Fahrkartenverkauf* (B), *Informationen* (C), *Komfort* (D) und *Sicherheit* (E) bezeichnet wurden. Die Objektpunkte wurden aus Gründen der Übersichtlichkeit unter Berücksichtigung der Cluster fortlaufend nummeriert, die zugehörigen Bezeichnungen der Qualitätsmerkmale befinden sich in Tab. 6.4.

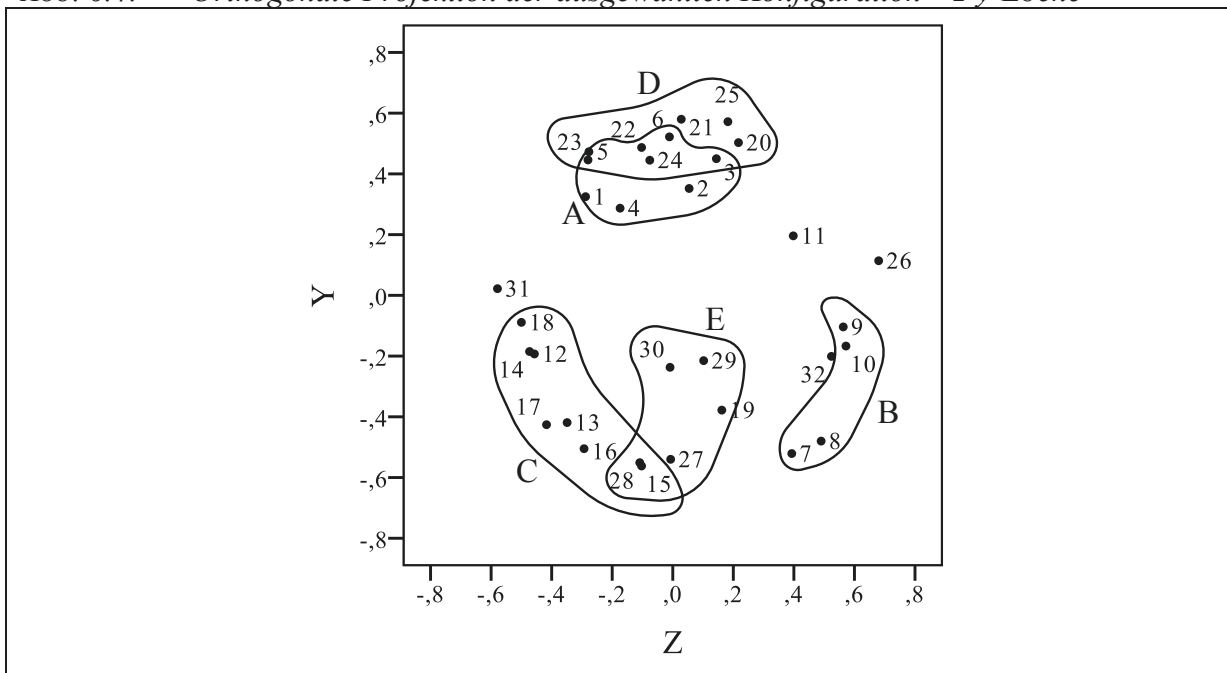
Abb. 6.2: Orthogonale Projektion der ausgewählten Konfiguration – x-y-Ebene



Quelle: Eigene Darstellung.

Abb. 6.3: Orthogonale Projektion der ausgewählten Konfiguration – x - z -Ebene

Quelle: Eigene Darstellung.

Abb. 6.4: Orthogonale Projektion der ausgewählten Konfiguration – z - y -Ebene

Quelle: Eigene Darstellung.

Eine alternative Interpretation besteht darin, die Cluster *Fahrkartenverkauf* (B) und *Sicherheit* (E) in jeweils zwei Cluster aufzuteilen, sodass sich insgesamt sieben Cluster ergeben. Im Cluster *Fahrkartenverkauf* (B) grenzen sich die Qualitätsmerkmale *Fahrkartenautomat am Bahnsteig* (7) und *Fahrkartenverkauf durch Personal am Bahnhof* (8) leicht von

Fahrkartenautomat im Zug (9) und Fahrkartenverkauf durch Personal im Zug (10) ab. Analog grenzen sich im Cluster Sicherheit (E) Sicherheitspersonal am Bahnsteig (27) und Videoüberwachung am Bahnsteig (28)

Tab. 6.4: Cluster und Qualitätsmerkmale

| A | Verbindungsqualität |
|----------|-------------------------------------|
| 1 | Anschlussgarantie |
| 2 | Frequenz |
| 3 | Gesamtreisezeit (von Tür zu Tür) |
| 4 | kein Fahrtenausfall |
| 5 | Pünktlichkeit |
| 6 | Umsteigehäufigkeit |

| B | Fahrkartenverkauf |
|----------|--|
| 7 | Fahrkartenautomat am Bahnsteig |
| 8 | Fahrkartenverkauf durch Personal am Bahnhof |
| 9 | Fahrkartenautomat im Zug |
| 10 | Fahrkartenverkauf durch Personal im Zug |

| | |
|----|------------------------------|
| 11 | Freundlichkeit des Personals |
|----|------------------------------|

| C | Informationen |
|----------|--|
| 12 | Durchsage im Zug zu Anschlusszügen |
| 13 | Durchsagen am Bahnsteig bei Störungen oder Verspätungen |
| 14 | Durchsagen im Zug bei Störungen oder Verspätungen |
| 15 | Fahrplaninformationen am Bahnsteig |
| 16 | Informationsanzeige am Bahnsteig bei Störungen oder Verspätungen |
| 17 | Informationsanzeige im Zug zu Anschlusszügen |
| 18 | Informationsanzeige am Zug (außen) |

| | |
|----|------------------------------|
| 19 | Fahrplanauskunft im Internet |
|----|------------------------------|

| D | Komfort |
|----------|--------------------------------|
| 20 | Abstellmöglichkeiten im Zug |
| 21 | Komfort der Sitzplätze |
| 22 | Sauberkeit der Toilette im Zug |
| 23 | Sauberkeit Zug außen |
| 24 | Sauberkeit Zug innen |
| 25 | Sitzplatzverfügbarkeit im Zug |

| | |
|----|-----------|
| 26 | Fahrpreis |
|----|-----------|

| E | Sicherheit |
|----------|-------------------------------------|
| 27 | Sicherheitspersonal am Bahnsteig |
| 28 | Videoüberwachung am Bahnsteig |
| 29 | Sicherheitspersonal im Zug |
| 30 | Videoüberwachung im Zug |

| | |
|----|---|
| 31 | Unterstellplatz/Wartehaus am Bahnsteig |
|----|---|

| | |
|----|--------------------------------------|
| 32 | Verständlichkeit des Tarifsystems |
|----|--------------------------------------|

Quelle: Eigene Darstellung.

leicht von *Sicherheitspersonal im Zug* (29) und *Videoüberwachung im Zug* (30) ab. Diese Abgrenzung resultiert aus den paradigmatischen Sortierungen durch die Befragten, d. h. die Befragten führten die Sortierung nach inhaltlicher Ähnlichkeit anhand unterschiedlicher Kriterien durch. So sortierten beispielsweise 84 % der Befragten die beiden Qualitätsmerkmale mit Bezug zur Sicherheit am Bahnsteig (27 und 28) in eine Gruppe, aber nur zwischen 61 % und 70 % sortierten mindestens eines der beiden Qualitätsmerkmale mit Bezug zur Sicherheit im Zug (29 und 30) zu einem Qualitätsmerkmal mit Bezug zur Sicherheit am Bahnsteig. Daraus wurde gefolgert, dass ein (geringer) Teil der Befragten eine Unterscheidung nach räumlichen Kriterien (im Zug/am Bahnsteig) vornahm.

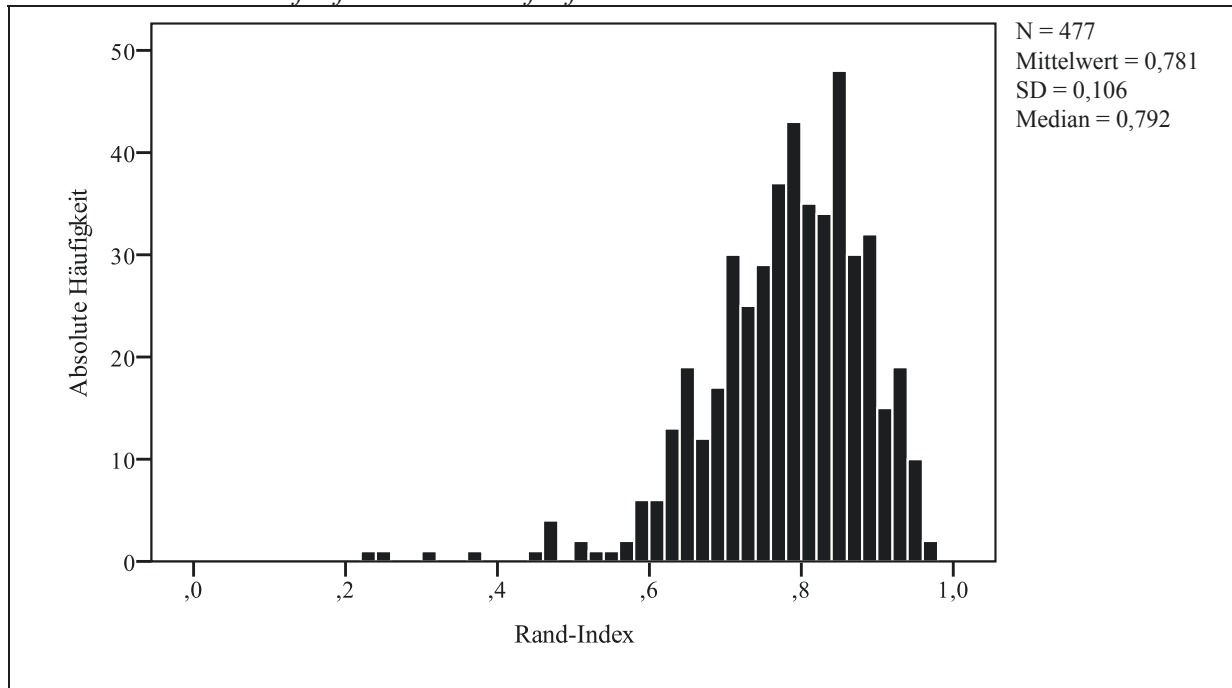
Sowohl bei der Interpretation mit fünf Clustern als auch bei der Interpretation mit sieben Clustern wurden fünf Qualitätsmerkmale identifiziert, die keinem Cluster zugeordnet werden konnten.

In der ermittelten Konfiguration wurden die aggregierten Unähnlichkeitsdaten bestmöglich wiedergegeben und die gefundenen fünf bzw. sieben Cluster können grundsätzlich als Konstrukte verwendet werden. Allerdings geht wie in Kap. 5.2 beschrieben eine gute Anpassung der Konfiguration an die aggregierten Daten nicht notwendigerweise mit einer guten Übereinstimmung zwischen der (interpretierten) Konfiguration und den Sorting-Daten auf Individualebene einher.

Die beiden konkurrierenden Interpretationen wurden als Partitionen angesehen und die Güte der Übereinstimmung mithilfe des Rand-Indexes ermittelt (vgl. Kap. 5.3). Dabei wurde jede einzelne der 477 Partitionen der Sorting-Daten auf Individualebene zum einen mit der ersten Partition (interpretierte Konfiguration mit fünf Clustern und fünf einzelnen Qualitätsmerkmalen) und zum anderen mit der zweiten Partition (interpretierte Konfiguration mit sieben Clustern und fünf einzelnen Qualitätsmerkmalen) verglichen, sodass insgesamt 954 Rand-Indizes bestimmt wurden.

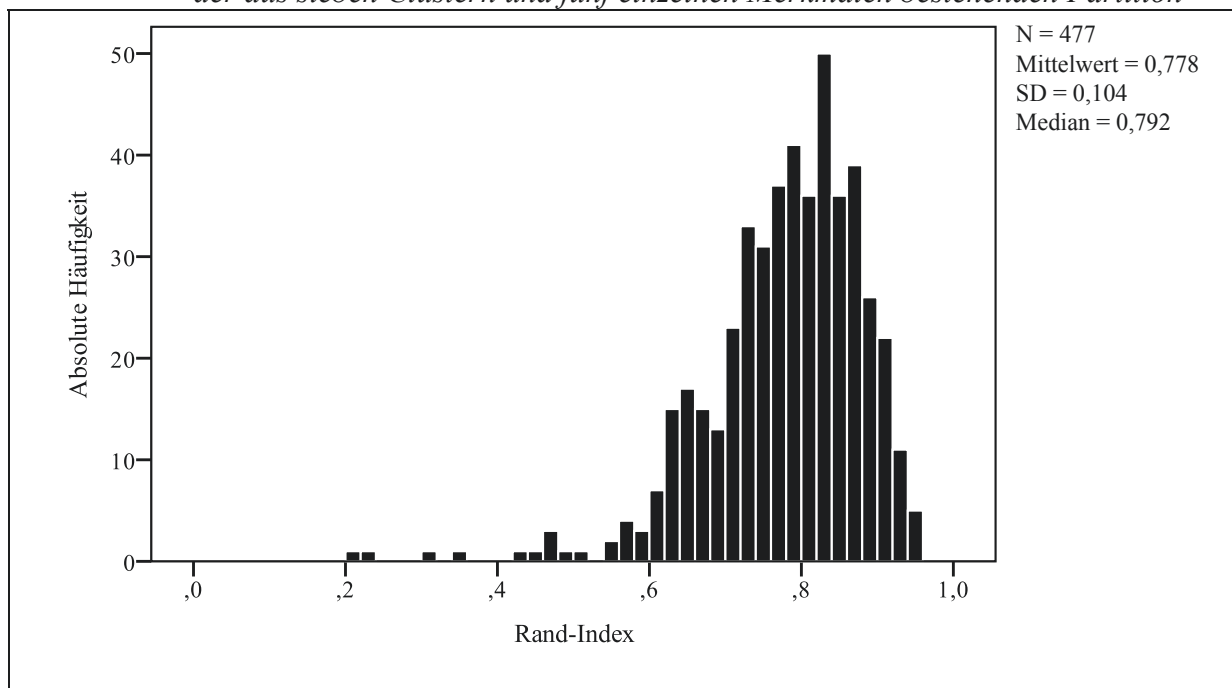
Die Verteilungen der Rand-Indizes sind differenziert nach den beiden konkurrierenden Partitionen in Abb. 6.5 und Abb. 6.6 dargestellt. Die erste Verteilung weist einen Mittelwert der Rand-Indizes von 0,781 (SD: 0,11) und die zweite Verteilung von 0,778 (SD: 0,10) auf. Der Median beider

Abb. 6.5: Verteilung der Rand-Indizes – Vergleich der Sorting-Daten der Bahnnutzer mit der aus fünf Clustern und fünf einzelnen Merkmalen bestehenden Partition



Quelle: Eigene Darstellung.

Abb. 6.6: Verteilung der Rand-Indizes – Vergleich der Sorting-Daten der Bahnnutzer mit der aus sieben Clustern und fünf einzelnen Merkmalen bestehenden Partition



Quelle: Eigene Darstellung.

Verteilungen der Rand-Indizes beträgt 0,792, jedoch unterscheiden sich die Rand-Indizes der beiden Partitionen signifikant voneinander.³⁶⁸ Größere Rand-Indizes weist die Partition mit fünf Clustern und fünf einzelnen Qualitätsmerkmalen auf, sodass ebendiese Partition ausgewählt wurde. Die ermittelten fünf Cluster stellen mögliche Konstrukte für die zu entwickelnden Verkehrsmittelwahlexperimente dar.

Bei der bisherigen Analyse wurden ausschließlich die Sorting-Daten der *Bahnnutzer* berücksichtigt. Allerdings sollten die hergeleiteten Konstrukte gleichermaßen für *Bahnnutzer* und *Pkw-Nutzer* gelten. Daher wurde die Übereinstimmung zwischen den 32 Partitionen der *Pkw-Nutzer* mit der Partition mit den fünf Clustern und fünf einzelnen Qualitätsmerkmalen ebenfalls anhand der Rand-Indizes überprüft.

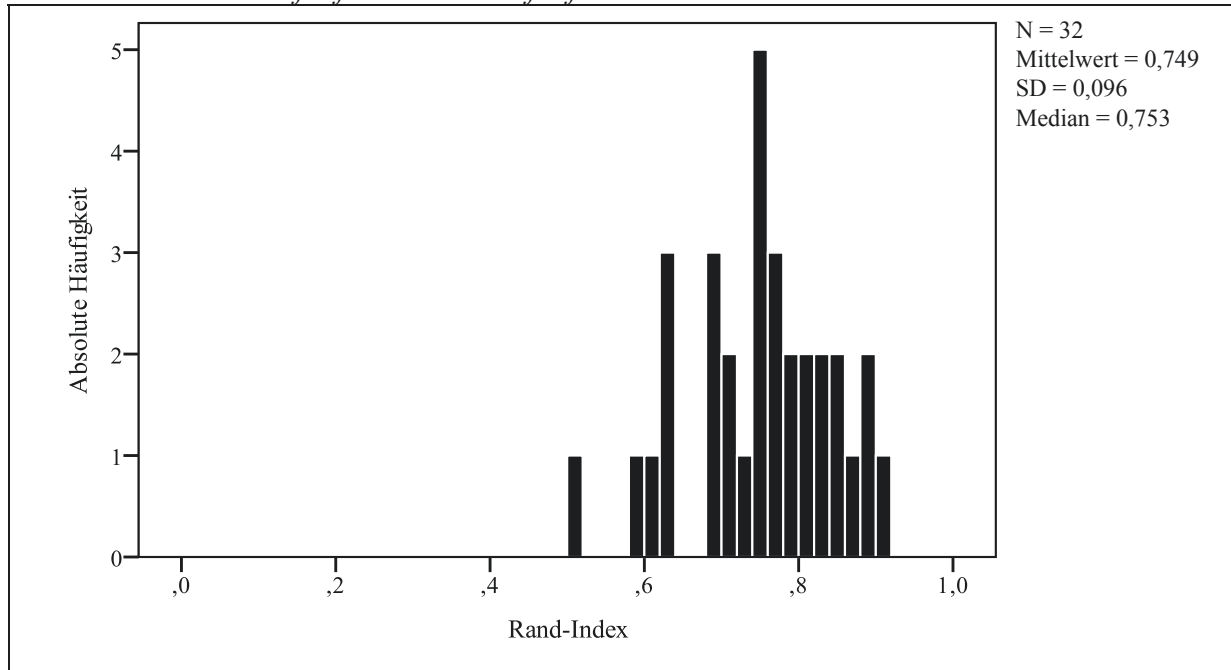
Die Verteilung dieser Rand-Indizes ist in Abb. 6.7 dargestellt. Der Mittelwert der Rand-Indizes beträgt 0,749 (SD: 0,10) und der Median 0,753. Die Rand-Indizes der Partitionen der *Pkw-Nutzer* unterscheiden sich nicht signifikant von denen der *Bahnnutzer*.³⁶⁹ Somit besteht auch eine gute Übereinstimmung der Partition mit fünf Clustern und fünf einzelnen Qualitätsmerkmalen mit den Partitionen der *Pkw-Nutzer*, welche nicht zur Bestimmung der MDS-Konfiguration herangezogen wurden. Daraus

³⁶⁸ Die Annahme der Gleichheit der Rand-Indizes der beiden Partitionen wurde mit dem Wilcoxon-Test (Vorzeichenrangtest), einem verteilungsfreien Test für zwei abhängige Stichproben, analysiert. Bei diesem Test wird für jedes Wertepaar die Differenz bestimmt und den absoluten Beträgen dieser Differenzen werden Rangplätze zugeordnet. Anschließend wird die Summe der Rangplätze der positiven Differenzen sowie die Summe der Rangplätze der negativen Differenzen berechnet. Die Nullhypothese, dass zwischen den beiden Stichproben keine systematischen Unterschiede bestehen, wird durch den Vergleich der Teststatistik (kleinere Summe der Rangplätze, die um die Anzahl aller Ränge und der Bindungen korrigiert wird) mit einem näherungsweise normalverteilten kritischen Wert überprüft, vgl. *Sedlmeier/Renkewitz* (2008), S. 589–594, *Wilcoxon* (1945), S. 80–81, *Bortz/Lienert* (2008), S. 191–200 und *Bühl* (2009), S. 354–358. In der vorliegenden Anwendung wurde die Nullhypothese der Gleichheit der Rand-Indizes der beiden Partitionen mit einem Signifikanz-Wert von $p=0,000$ abgelehnt.

³⁶⁹ Die Annahme der Gleichheit der Rand-Indizes der Partitionen der *Pkw-* und *Bahnnutzer* wurde mit dem Mediantest, einem verteilungsfreien Parametertest für zwei unabhängige und ungleich große Stichproben, analysiert. Beim Mediantest wird für beide Stichproben ein gemeinsamer Median ermittelt und anschließend gezählt, wie viele Werte in beiden Stichproben jeweils oberhalb und unterhalb dieses Medians liegen. Die Vierfeldertafel wird eingesetzt, um die Nullhypothese, dass sich die beiden Stichproben hinsichtlich des Medians nicht unterscheiden, anhand eines Chi-Quadrat-Tests zu überprüfen, vgl. *Bortz/Lienert* (2008), S. 137–142 und *Bühl* (2009), S. 363–364. Die Nullhypothese der Gleichheit der Mediane wurde mit einem Signifikanzwert von $p=0,108$ nicht abgelehnt.

wurde gefolgert, dass die möglichen Konstrukte für die zu entwickelnden Verkehrsmittelwahlexperimente gleichermaßen bei Befragungen von *Pkw-Nutzern* eingesetzt werden können.

Abb. 6.7: Verteilung der Rand-Indizes – Vergleich der Sorting-Daten der Pkw-Nutzer mit der aus fünf Clustern und fünf einzelnen Merkmalen bestehenden Partition



Quelle: Eigene Darstellung.

6.5 Analyse der Ratingdaten

Die Analyse der Ratingdaten erfolgte differenziert nach den beiden Nutzergruppen mit Methoden der deskriptiven Statistik. Eine Übersicht über ausgewählte Lage- und Streuungsmaße ist in Tab. 6.5 dargestellt.

Allgemein besteht bei Rating-Experimenten das Risiko einer *Anspruchsinflation*³⁷⁰, d. h. Befragte neigen dazu, alle Merkmale als (sehr) wichtig zu bewerten. Auch bei den im Rahmen dieser Vorstudie erhobenen Ratingdaten ist insgesamt eine Tendenz zu hohen Bewertungen zu erkennen: Die beiden höchsten Bewertungen „6“ bzw. „7“ wurden bei 51,9 % aller Bewertungen (51,4 % der Bewertungen der *Bahnnutzer* und

³⁷⁰ Eine solche *Anspruchsinflation* ist möglich, weil der Befragte nicht hinsichtlich der Anzahl der Höchstbewertungen eingeschränkt ist. Dieses Problem wird beispielsweise bei Experimenten mit Konstantsummenverfahren, mit einer Rangreihenbildung oder mit (diskreten) Wahlentscheidungen umgangen, vgl. *Matzler/Bailom* (2002), S. 235–236, *Helm/Steiner* (2008), S. 248–249, *Berekoven/Eckert/Ellenrieder* (2004), S. 78, *Kroeber-Riel/Weinberg/Gröppel-Klein* (2009), S. 221 und *Matzler/Sauerwein/Stark* (2002), S. 282.

Tab. 6.5: Lage- und Streuungsmaße der Ratingdaten differenziert nach Bahnnutzern und Pkw-Nutzern

| | Bahnnutzer | | | | | Pkw-Nutzer | | | | |
|---|------------|------------|------|--------|-----|------------|------------|------|--------|-----|
| | N | Mittelwert | SD | Median | IQR | N | Mittelwert | SD | Median | IQR |
| 1 Anschlussgarantie | 434 | 5,81 | 1,39 | 6 | 2 | 30 | 6,27 | 1,28 | 7 | 1 |
| 2 Frequenz | 433 | 5,38 | 1,18 | 5 | 1 | 30 | 5,50 | 0,97 | 5 | 1 |
| 3 Gesamtreisezeit | 434 | 5,21 | 1,22 | 5 | 2 | 30 | 5,13 | 1,07 | 5 | 1 |
| 4 Kein Fahrtenausfall | 433 | 6,45 | 0,86 | 7 | 1 | 30 | 6,30 | 1,15 | 7 | 1 |
| 5 Pünktlichkeit | 434 | 6,24 | 0,94 | 6 | 1 | 30 | 6,27 | 0,78 | 6 | 1 |
| 6 Umsteigehäufigkeit | 433 | 4,91 | 1,59 | 5 | 2 | 30 | 5,33 | 1,56 | 6 | 3 |
| 7 Fahrkartenautomat am Bahnsteig | 415 | 5,04 | 1,82 | 5 | 3 | 29 | 5,28 | 1,39 | 5 | 1 |
| 8 Fahrkartenverkauf durch Personal am Bahnhof | 403 | 4,46 | 1,82 | 5 | 3 | 29 | 5,03 | 1,95 | 5 | 3 |
| 9 Fahrkartenautomat im Zug | 414 | 4,85 | 1,80 | 5 | 2 | 29 | 5,14 | 1,33 | 5 | 1 |
| 10 Fahrkartenverkauf durch Personal im Zug | 411 | 4,50 | 1,79 | 5 | 3 | 29 | 5,31 | 1,49 | 6 | 3 |
| 11 Freundlichkeit des Personals | 416 | 5,85 | 1,23 | 6 | 2 | 29 | 6,07 | 1,16 | 7 | 2 |
| 12 Durchsage im Zug zu Anschlusszügen | 418 | 4,75 | 1,64 | 5 | 2 | 30 | 5,47 | 1,53 | 6 | 2 |
| 13 Durchsagen am Bahnsteig bei Störungen oder Verspätungen | 427 | 6,37 | 0,98 | 7 | 1 | 30 | 6,47 | 0,68 | 7 | 1 |
| 14 Durchsagen im Zug bei Störungen oder Verspätungen | 427 | 6,31 | 1,01 | 7 | 1 | 30 | 6,47 | 0,73 | 7 | 1 |
| 15 Fahrplaninformationen am Bahnsteig | 429 | 5,70 | 1,36 | 6 | 2 | 30 | 5,97 | 0,89 | 6 | 2 |
| 16 Informationsanzeige am Bahnsteig bei Störungen oder Verspätungen | 425 | 6,26 | 1,02 | 7 | 1 | 30 | 6,43 | 0,63 | 7 | 1 |
| 17 Informationsanzeige im Zug zu Anschlusszügen | 409 | 4,82 | 1,56 | 5 | 2 | 30 | 5,17 | 1,51 | 5 | 3 |
| 18 Informationsanzeige am Zug (außen) | 419 | 4,64 | 1,63 | 5 | 2 | 30 | 5,57 | 1,14 | 6 | 1 |
| 19 Fahrplanauskunft im Internet | 428 | 5,81 | 1,42 | 6 | 2 | 30 | 6,17 | 0,87 | 6 | 1 |
| 20 Abstellmöglichkeiten im Zug | 432 | 4,12 | 1,80 | 4 | 3 | 30 | 4,37 | 1,73 | 5 | 3 |
| 21 Komfort der Sitzplätze | 433 | 4,43 | 1,46 | 4 | 2 | 30 | 4,13 | 1,50 | 5 | 2 |

Fortsetzung nächste Seite

Fortsetzung Tab. 6.5

| | Bahnnutzer | | | | | Pkw-Nutzer | | | | |
|---|------------|------------|------|--------|-----|------------|------------|------|--------|-----|
| | N | Mittelwert | SD | Median | IQR | N | Mittelwert | SD | Median | IQR |
| 22 Sauberkeit der Toilette im Zug | 430 | 5,40 | 1,69 | 6 | 3 | 30 | 5,97 | 1,65 | 7 | 1 |
| 23 Sauberkeit Zug außen | 429 | 3,39 | 1,66 | 3 | 2 | 30 | 3,17 | 1,49 | 3 | 2 |
| 24 Sauberkeit Zug innen | 430 | 5,63 | 1,17 | 6 | 2 | 30 | 6,03 | 0,96 | 6 | 2 |
| 25 Sitzplatzverfügbarkeit im Zug | 433 | 4,94 | 1,51 | 5 | 2 | 30 | 4,9 | 1,52 | 5 | 2 |
| 26 Fahrpreis | 432 | 5,65 | 1,40 | 6 | 2 | 30 | 5,80 | 1,24 | 6 | 2 |
| 27 Sicherheitspersonal am Bahnsteig | 417 | 5,13 | 1,68 | 6 | 3 | 30 | 5,67 | 1,45 | 6 | 2 |
| 28 Videoüberwachung am Bahnsteig | 417 | 5,06 | 1,79 | 5 | 3 | 30 | 5,47 | 1,93 | 6 | 2 |
| 29 Sicherheitspersonal im Zug | 416 | 4,96 | 1,70 | 5 | 2 | 30 | 5,03 | 2,09 | 6 | 3 |
| 30 Videoüberwachung im Zug | 417 | 4,77 | 1,83 | 5 | 2 | 30 | 5,17 | 1,95 | 6 | 3 |
| 31 Unterstellplatz / Wartehaus am Bahnsteig | 419 | 4,17 | 1,70 | 4 | 2 | 30 | 4,97 | 1,45 | 5 | 2 |
| 32 Verständlichkeit des Tarifsystems | 411 | 6,00 | 1,23 | 6 | 2 | 29 | 6,62 | 0,62 | 7 | 1 |

Quelle: Eigene Erhebung.

58,8 % der Bewertungen der *Pkw-Nutzer*) vergeben. Die höchste Bewertung „7“ erfolgte in 28,8 % aller Bewertungen (28,5 % der Bewertungen der *Bahnnutzer* und 33,4 % der Bewertungen der *Pkw-Nutzer*). Niedrige Bewertungen wurden hingegen eher selten vorgenommen: Nur bei 8,3 % aller Bewertungen wurde eine „1“ oder „2“ vergeben (8,5 % der Bewertungen der *Bahnnutzer* und 5,7 % der *Pkw-Nutzer*). Eine Tendenz zur Mitte³⁷¹, die bei Rating-Skalen mit einer ungeraden Anzahl von Stufe auftreten kann, ist nicht zu erkennen. Die mittlere Stufe „4“ wurde nur bei 12,7 % aller Bewertungen (13,0 % der Bewertungen der *Bahnnutzer*, 9,0 % der Bewertungen der *Pkw-Nutzer*) verwendet.

Die Bewertungen der einzelnen Merkmale unterscheiden sich hinsichtlich der Mediane. Bei beiden Nutzergruppen reicht der Median jeweils von 3 bis 7. Der Interquartilsabstand IQR (definiert als Differenz der Quartile

³⁷¹ Vgl. Bortz/Döring (2006), S. 184 und Jankisz/Moosbrugger (2007), S. 60–61.

Q_{.25} und Q_{.75}) streut jeweils zwischen 1 und 3, wobei ein höherer Interquartilsabstand auf eine höhere Heterogenität der Bewertungen hindeutet.

Konservativ wird bei Ratingdaten ein ordinales Skalenniveau angenommen.³⁷² Für die Bildung des Mittelwerts und der Standardabweichung ist jedoch ein metrisches Skalenniveau erforderlich, welches für eine bessere Differenzierung der Merkmale anhand der entsprechenden Kennzahlen unterstellt wurde. Sowohl bei den Daten der *Bahnnutzer* als auch bei den *Pkw-Nutzern* handelt es sich bei *Sauberkeit Zug außen* (23) mit 3,39 bzw. 3,17 um das Merkmal mit dem niedrigsten Mittelwert. Den höchsten Mittelwert weist bei den *Bahnnutzern* das Merkmal *kein Fahrtenausfall* (4) mit 6,45 und bei den *Pkw-Nutzern* *Verständlichkeit des Tarifsystems* (32) mit 6,62 auf.

Die höchste bzw. niedrigste Standardabweichung der Bewertungen der *Bahnnutzer* weisen die Merkmale *Videoüberwachung im Zug* (30) mit 1,83 bzw. *kein Fahrtenausfall* (4) mit 0,86 auf. Insbesondere die Merkmale mit Bezug zum Fahrkartenverkauf und zur Sicherheit wurden sehr heterogen bewertet. Auch von den *Pkw-Nutzern* wurden insbesondere die Merkmale mit Bezug zur Sicherheit sehr heterogen bewertet. Das Merkmal *Sicherheitspersonal im Zug* (29) weist mit 2,09 die höchste Standardabweichung auf; bei *Verständlichkeit des Tarifsystems* (32) handelt es sich mit 0,62 um das Merkmal mit der geringsten Standardabweichung.

Anhand der erhobenen Daten kann nicht analysiert werden, warum die Merkmalsbewertung mit Bezug zur Sicherheit relativ große Standardabweichungen aufweisen. Allerdings kann vermutet werden, dass die teilweise geringen Bewertungen mit einer kritischen Einstellung gegenüber einer Videoüberwachung einhergehen. In Akzeptanzstudien zur Videoüberwachung wird diese teilweise kritische Einstellung deutlich.³⁷³

Hinsichtlich der Bewertung der Merkmale mit Bezug zum Fahrkartenverkauf sei an dieser Stelle erneut auf den relativ hohen Anteil der Befragten, die über ein Monatsticket verfügen, hingewiesen.

³⁷² Vgl. Bortz/Döring (2006), S. 181–182, Jankisz/Moosbrugger (2007), S. 52, Sedlmeier/Renkewitz (2008), S. 66 und Bühner (2006), S. 331–332.

³⁷³ Vgl. dazu beispielsweise Helten/Fischer (2004) und Hempel/Töpfer (2004).

6.6 Auswahl von Konstrukten und Attributen

Nicht alle Qualitätsmerkmale und Cluster konnten als Attribute und Konstrukte in die Verkehrsmittelwahlexperimente übernommen werden. Denn trotz einer Zusammenfassung der Attribute zu Konstrukten wären die hierarchischen Experimente und die vorangestellten Rating-Experimente derart komplex und zeitintensiv, dass die dabei vorzunehmende Datenerhebung nicht praktikabel erschien.

Ein Attribut, das Teil eines Konstrukts ist, wird nur in einem Subexperiment direkt als Attribut verwendet; in die anderen Subexperimente wird es nur indirekt durch den zusammenfassenden Konstruktwert einbezogen. Dahingegen müsste ein Attribut, das nicht Teil eines Konstrukts ist, in jedes Subexperiment als Attribut einbezogen werden. Die Anzahl der dem Befragten gleichzeitig vorgelegten Attribute würde sich dadurch entsprechend erhöhen. Aus diesem Grund wurden die Merkmale ohne Zuordnung zu einem Cluster weitestgehend aus der Untersuchung ausgeschlossen. Dazu zählen die Qualitätsmerkmale *Freundlichkeit des Personals* (11), *Fahrplanauskunft im Internet* (19), *Unterstellplatz/Wartehaus am Bahnsteig* (31) und *Verständlichkeit des Tarifsystems* (32).³⁷⁴ Eine Ausnahme bei den Merkmalen ohne Zuordnung zu einem Cluster bildet der *Fahrpreis* (26). Dieses Merkmal wurde als einzelnes Attribut in alle Subexperimente einbezogen um eine Bestimmung von Zahlungsbereitschaften für alle Attributsausprägungen und zusammenfassende Konstruktwerte zu ermöglichen.

Zudem konnten nicht alle der fünf Cluster als Konstrukte einbezogen werden. Cluster, deren Merkmalsbewertungen relativ niedrige Mediane bzw. Mittelwerte und/oder hohe Interquartilsabstände bzw. Standardabweichungen aufweisen, wurden aus der weiteren Untersuchung ausgeschlossen: Dazu zählen die Merkmale des Clusters *Fahrkartenverkauf* und die Merkmale des Clusters *Sicherheit*, sodass diese beiden Cluster nicht als Konstrukte übernommen wurden.

³⁷⁴ Aufgrund der relativ hohen und homogenen Bewertungen des Merkmals *Verständlichkeit des Tarifsystems* seitens der *Pkw-Nutzer* besteht hier ein Ansatzpunkt für weitere Untersuchungen außerhalb der vorliegenden Arbeit.

Die drei Cluster *Verbindungsqualität*, *Komfort* und *Informationen* wurden als Konstrukte übernommen, wobei jedoch nicht alle der zugehörigen Merkmale als Attribute berücksichtigt werden konnten. Die Merkmale *Sauberkeit Zug außen* (23), *Informationsanzeige am Zug (außen)* (17) und *Abstellmöglichkeiten im Zug (z. B. für Gepäck, Fahrräder, Kinderwagen)* (20) wurden aufgrund relativ niedriger und/oder heterogenerer Bewertungen aus der weiteren Untersuchung ausgeschlossen. Die verbleibenden Merkmale wurden als Attribute übernommen, wobei jedoch in manchen Fällen zwei Merkmale zu einem Attribut zusammengefasst wurden.

Bei drei zu analysierenden Konstrukten mussten drei Subexperimente gebildet werden. Um für alle drei Subexperimente die Berechnung der Zahlungsbereitschaft für eine Veränderung der Gesamtreisezeit (Wert der Zeit) zu ermöglichen, wurde das Merkmal *Gesamtreisezeit (von Tür zu Tür)* (3) aus dem Konstrukt der *Verbindungsqualität* herausgenommen und als einzelnes Attribut in allen Subexperimenten berücksichtigt.

6.7 Zusammenfassung

Zur Entwicklung der Konstrukte für die hierarchischen Verkehrsmittelwahllexperimente wurde eine empirische Vorstudie durchgeführt. In Befragungen von *Bahnnutzern*, aber auch von *Pkw-Nutzern* (Kontrollgruppe), wurden für 32 Qualitätsmerkmale (Un-)Ähnlichkeitsdaten und Ratingdaten erhoben. Die Unähnlichkeitsdaten der *Bahnnutzer* wurden in aggregierter Form mit einem Verfahren der multidimensionalen Skalierung für ratioskalierte Daten mit dem *Proxscal*-Algorithmus ausgewertet. Vor dem Hintergrund des Anpassungsmaßes und der Interpretierbarkeit wurde aus mehreren Konfigurationen unterschiedlicher Dimensionalität die Konfiguration mit drei Dimensionen ausgewählt. In dieser Konfiguration wurden fünf Cluster identifiziert, von denen sich zwei Cluster möglicherweise in jeweils zwei Cluster aufteilen. Fünf Merkmale konnten keinem Cluster zugeordnet werden.

Bei diesen beiden konkurrierenden Interpretationen handelt es sich näherungsweise um Partitionen. Zur Überprüfung der Übereinstimmung zwischen den interpretierten Konfigurationen und den Sorting-Daten auf Individualebene sowie zur Auswahl zwischen den konkurrierenden Inter-

pretationen wurden Rand-Indizes gebildet. Auf Basis der Verteilungen der Rand-Indizes wurde die Partition mit fünf Clustern ausgewählt. Zudem wurde die Übereinstimmung dieser Partition mit den Sorting-Daten der *Pkw-Nutzern* überprüft und kein signifikanter Unterschied festgestellt. Daraus wurde gefolgert, dass die Cluster grundsätzlich als Konstrukte in hierarchischen Wahlexperimenten gleichermaßen für *Bahnnutzer* als auch für *Pkw-Nutzer* geeignet sind.

Allerdings konnten nicht alle Qualitätsmerkmale und Cluster als Attribute und Konstrukte in die Verkehrsmittelwahlexperimente übernommen werden, ohne dass diese für die geplante Datenerhebung zu komplex bzw. zeitintensiv geworden wären. Eine Auswahl der Attribute und Konstrukte erfolgte vor dem Hintergrund der Bewertung der Wichtigkeit der Qualitätsmerkmale. Im Ergebnis wurden die drei Cluster *Verbindungsqualität*, *Komfort* und *Informationen* als Konstrukte übernommen, wobei nicht alle enthaltenen Merkmale als Attribute berücksichtigt werden konnten. Die Merkmale *Fahrpreis* und *Gesamtreisezeit* wurden als einzelne Attribute in allen Subexperimenten berücksichtigt, um die Berechnung von Zahlungsbereitschaften, insbesondere für eine Veränderung der Gesamtreisezeit (Wert der Zeit), zu ermöglichen.

7 Anwendung des Verfahrens der HII-I im Regionalverkehr

Unter Verwendung der drei in der Vorstudie hergeleiteten Konstrukte und deren Attribute (vgl. Kap. 6) wird im Folgenden das Verfahren der HII-I im Regionalverkehr angewendet. Dazu wird ein Erhebungsinstrument entwickelt, entsprechende Befragungen von Nutzern und Nicht-Nutzern des regionalen SPNV werden durchgeführt und im Anschluss daran die Daten analysiert.

7.1 Untersuchungsdesign und Datenerhebung

Ebenso wie die Vorstudie wurde die Hauptstudie als computergestützte Face-to-face-Befragung konzipiert. Dieses Erhebungsverfahren wurde ausgewählt, da durch die Verwendung eines Laptops eine individuen-spezifische Wahlsituation mit Bezug zu einer konkreten Fahrt während der Befragung generiert werden konnte und durch die persönliche Befragung vergleichsweise längere und komplexere Fragestellungen möglich waren. In Analogie zur Vorstudie wurden als Erhebungsorte zur Befragung der Nutzer und Nicht-Nutzer des regionalen SPNV die Züge der Kooperationspartner bzw. Behörden ausgewählt.

Zur Entwicklung des Untersuchungsdesigns war es erforderlich, für die in Kap. 6.6 ausgewählten Attribute und Konstrukte zur Charakterisierung der Verkehrsmittel geeignete Ausprägungen zu bestimmen. Die Attributsausprägungen wurden in Absprache mit den Kooperationspartnern zum einen im Sinne einer möglichst großen praktischen Relevanz ausgewählt, beispielsweise aufgrund von Vorgaben in Verkehrsdienstleistungsverträgen; zum anderen musste jedoch im Sinne der Verständlichkeit bzw. Vorstellbarkeit auf zu technisch definierte Ausprägungen verzichtet und stattdessen eher subjektiv interpretierbare Ausprägungen³⁷⁵ verwendet werden. Eine Übersicht der Ausprägungen der qualitativen Attribute ist in Tab. 7.1 dargestellt.

³⁷⁵ Diese Problematik bestand insbesondere bei den Ausprägungen der Attribute des *Komforts*.

Bei den Ausprägungen des Attributs *Takt*³⁷⁶ erschien eine Orientierung an den aktuellen Fahrplänen sinnvoll. Im Untersuchungsgebiet führen die meisten Züge der ausgewählten Linien tagsüber im Stundentakt, zu morgendlichen und abendlichen Stoßzeiten im Halbstundentakt und zu Tagesrandzeiten im Zweistundentakt. Die gewählten Ausprägungen mit 30, 60 und 120 Minuten erschienen daher realistisch.³⁷⁷ Dem Attribut *Pünktlichkeit* wurden die Ausprägungen *pünktlich bis max. 3 Minuten Verspätung*, *3 bis 10 Minuten Verspätung* und *10 bis 20 Minuten Verspätung* zugewiesen. Bei einem Halbstundentakt kommt eine Verspätung von 20 Minuten annähernd einem Fahrtenausfall gleich, sodass ein Fahrtenausfall nicht explizit im Experiment berücksichtigt wurde. Die auf den ersten Blick eher unüblich erscheinende Schwelle von 3 Minuten Verspätung wurde deshalb gewählt, weil gemäß den Verkehrsdienstleistungsverträgen eine Verzögerung um bis zu 3 Minuten noch als pünktlich gewertet wird; eine Schwelle von 5 Minuten könnte diesen Malus-relevanten Aspekt nicht berücksichtigen.³⁷⁸ In die Ausprägungen des Attributs *Umsteigehäufigkeit* wurde auch die Anschlussgarantie einbezogen.

Beim Attribut *Toilette im Zug* wurde abhängig vom Vorhandensein zwischen einer *sauberen* und einer *verschmutzten Toilette* unterschieden. Unabhängig von der *Sitzplatzverfügbarkeit im Zug* wurde beim *Komfort der Sitzplätze* zwischen *bequemen* und *unbequemen* Sitzplätzen unterschieden.

Bei den drei Attributen *Informationen im Zug zu Grund und Dauer von Störungen*, *Informationen am Bahnhof zu Grund und Dauer von Störungen* und *Informationen im Zug zu aktuellen Anschlusszügen* wurde unterschieden, ob und auf welche Weise Informationen bereitgestellt werden. Mit *Durchsage und Anzeige*, *Durchsage*, *Anzeige* und *keine Informationen* wurden mehrere Merkmale der Vorstudie als Ausprägungen eines Attributs zusammengefasst.

³⁷⁶ Da die Bezeichnung *Frequenz* in der Vorstudie häufig einer Erklärung bedurfte, wurde für die weitere Untersuchung stattdessen die Bezeichnung *Takt* verwendet.

³⁷⁷ Ein 20- oder 90-Minutentakt wäre aufgrund des integralen Fahrplans des bestehenden Netzes mit Problemen verbunden.

³⁷⁸ Vgl. *Nahverkehr Westfalen-Lippe* (2011), S. 85.

Tab. 7.1: Ausprägungen der qualitativen Attribute der Konstrukte Verbindungsqualität, Komfort und Informationen

| Verbindungsqualität | Komfort | Informationen |
|--|--|---|
| <p>Pünktlichkeit</p> <ol style="list-style-type: none"> 1. pünktlich bzw. max. 3 Minuten Verspätung 2. 3 bis 10 Minuten Verspätung 3. 10 bis 20 Minuten Verspätung <p>Umsteigehäufigkeit</p> <ol style="list-style-type: none"> 1. 0 x Umsteigen 2. 1 x Umsteigen mit Anschlussgarantie 3. 1 x Umsteigen ohne Anschlussgarantie <p>Takt</p> <ol style="list-style-type: none"> 1. alle 30 Minuten 2. alle 60 Minuten 3. alle 120 Minuten | <p>Toilette im Zug</p> <ol style="list-style-type: none"> 1. sauber 2. verschmutzt 3. nicht vorhanden <p>Sauberkeit Zug innen</p> <ol style="list-style-type: none"> 1. sauber 2. Boden klebrig, Papier liegt auf den Sitzen <p>Sitzplatzverfügbarkeit im Zug</p> <ol style="list-style-type: none"> 1. ganze Fahrt Sitzplatz 2. halbe Fahrt Sitzplatz 3. kein freier Sitzplatz <p>Komfort der Sitzplätze</p> <ol style="list-style-type: none"> 1. bequem 2. unbequem | <p>Fahrplaninformationen am Bahnhof</p> <ol style="list-style-type: none"> 1. vorhanden 2. nicht vorhanden <p>Informationen im Zug zu Grund und Dauer von Störungen</p> <ol style="list-style-type: none"> 1. Durchsage und Anzeige 2. Durchsage 3. Anzeige 4. keine Informationen <p>Informationen am Bahnhof zu Grund und Dauer von Störungen</p> <ol style="list-style-type: none"> 1. Durchsage und Anzeige 2. Durchsage 3. Anzeige 4. keine Informationen <p>Informationen im Zug zu aktuellen Anschlusszügen</p> <ol style="list-style-type: none"> 1. Durchsage und Anzeige 2. Durchsage 3. Anzeige 4. keine Informationen |

Quelle: Eigene Darstellung.

Zur Beschreibung des Regionalbusses wurden in den Bezeichnung der Attribute und deren Ausprägungen die Begriffe Zug und Bahnhof durch Bus und Bushaltestelle ersetzt.

Bei den Ausprägungen der Attribute *Gesamtreisezeit* und *Fahrpreis/Fahrtkosten* wurden absolute und relative Veränderungen ausgewählt, mit denen ausgehend von den Werten einer konkreten Fahrt während der Befragung Werte generiert wurden, um so einen Bezug zu dieser

konkreten Fahrt herzustellen.³⁷⁹ Beim Attribut *Gesamtreisezeit* wurden absolute Veränderungen von -10 Minuten, -5 Minuten, 0 Minuten, +5 Minuten und +10 Minuten und beim Attribut *Fahrpreis/Fahrtkosten* relative Veränderungen von 90 %, 95 %, 100 %, 105 % und 110 % (mit Rundung auf 10-Cent-Beträge) ausgewählt.

Als Ausprägungen der hypothetischen, zusammenfassenden Konstruktwerte wurden „+++“³⁸⁰, „++“ und „-“ (auf einer von „-----“ bis „++++“ reichenden Skala) festgelegt.³⁸⁰

Wie bereits bei der Vorstudie erfolgte die Programmierung des für die Datenerhebung erforderlichen Erhebungsinstruments mit *Visual Basic for Applications (VBA)* in *MS Access 2007*. Im Folgenden wird zunächst das Erhebungsinstrument für die Befragung der Nutzer des regionalen SPNV (Datenerhebung in den Zügen der Kooperationspartner) beschrieben, im Anschluss daran die vorgenommenen Änderungen des Erhebungsinstruments für die Befragung der Nicht-Nutzer des regionalen SPNV (Datenerhebung in den Behörden).

Dem Erhebungsinstrument für die Befragung der Nutzer des regionalen SPNV lag ein Fragebogen zugrunde, der sich in folgende Teile gliederte:

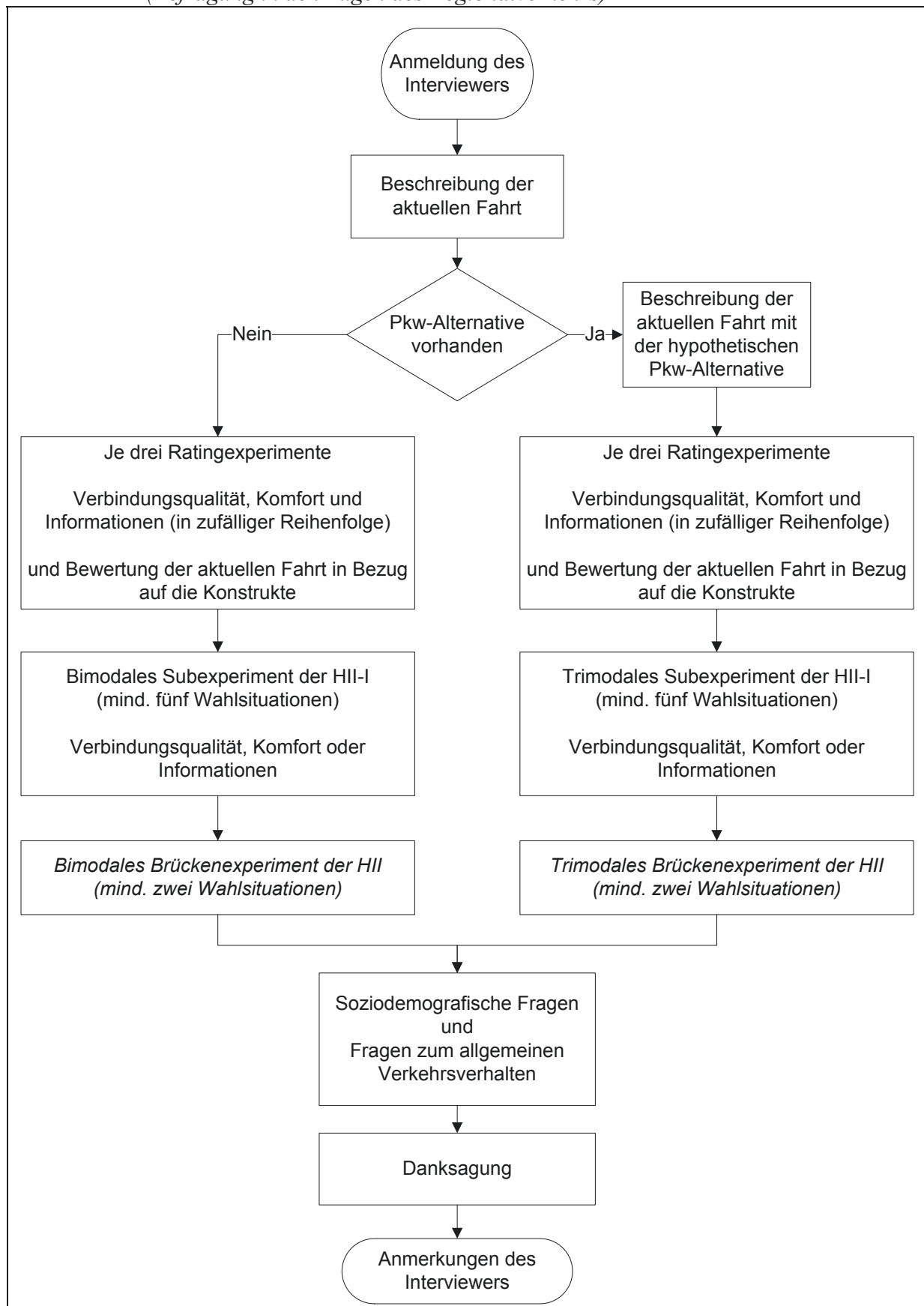
1. Anmeldung des Interviewers,
2. Beschreibung der aktuellen Fahrt,
3. Rating-Experimente und Bewertung der aktuellen Fahrt in Bezug auf die Konstrukte,
4. Subexperiment der HII-I,
5. Brückenexperiment der HII,³⁸¹
6. Soziodemografische Fragen und Fragen zum allgemeinen Verkehrsverhalten,
7. Danksagung und
8. Anmerkungen des Interviewers.

³⁷⁹ Attribute mit Ausprägungen, die relativ um einen Ausgangspunkt schwanken, werden als „pivoting attributes“ bezeichnet, vgl. *Hensher/Rose/Greene* (2005), S. 181–182.

³⁸⁰ Eine stark negative Bewertung, beispielsweise „-----“, wurde nicht ausgewählt. Diese Entscheidung beruhte auf den Ergebnissen eines Pretests, welcher zeigte, dass Befragte ihre aktuelle Fahrt in Bezug auf die drei Konstrukte eher positiv bewerteten.

³⁸¹ Auf die Anwendung der HII wird im Rahmen dieser Arbeit nicht weiter eingegangen.

Abb. 7.1: Aufbau des eingesetzten Erhebungsinstruments zur Befragung der Nutzer des regionalen SPNV (Befragung in den Zügen des Regionalverkehrs)



Quelle: Eigene Darstellung.

Der prinzipielle Aufbau des Erhebungsinstruments ist in Abb. 7.1 dargestellt. Eine Übersicht der einzelnen Fragebogenmasken befindet sich in Anhang D (Abb. D.1 bis Abb. D.13).

Nach der Anmeldung des Interviewers (Abb. D.1) wurde im Rahmen der Beschreibung der aktuellen Fahrt (Abb. D.2) zunächst der Startbahnhof und das zum Erreichen des Startbahnhofes eingesetzte Verkehrsmittel (inkl. „zu Fuß“) erfragt, sowie analog dazu der Zielbahnhof und das zum Erreichen des Ziels eingesetzte Verkehrsmittel. Daran anschließend wurde die Dauer der Fahrt insgesamt (für eine Richtung), die Anzahl der Umsteigevorgänge zwischen öffentlichen Verkehrsmitteln, der (Quelle-Ziel-bezogene) Fahrtzweck sowie die Fahrtkosten für diese Fahrt erfragt. Bei Verwendung einer Monatskarte wurde basierend auf dem Gesamtpreis und den durchschnittlich durchgeführten Fahrten gemeinsam mit dem Befragten ein Einzelpreis für die aktuelle Fahrt errechnet bzw. geschätzt.

Zudem wurde abgefragt, ob die Befragten die aktuelle Fahrt auch alternativ mit einem Pkw durchführen können. Sofern dieses alternative Verkehrsmittel zur Verfügung stand, wurden dafür die geschätzten Fahrtkosten und die Gesamtreisezeit erfragt.

Diese Fragen sollten zum einen dazu dienen, dass der Befragte die aktuelle Fahrt konkretisiert und sich bei seinen Antworten auf diese Fahrt bezieht. Zum anderen wurden ausgehend von diesen Daten die Ausprägungen für Gesamtreisezeit und Fahrtkosten der Pkw-Alternative – sofern diese Alternative vorhanden war – in den hierarchischen Wahlexperimenten absolut bzw. relativ variiert. Darüber hinaus bilden diese Daten eine Grundlage für eine Simulation von Wahlwahrscheinlichkeiten.

Im Anschluss an die Beschreibung der aktuellen Fahrt folgten die Rating-Experimente (Abb. D.3 bis Abb. D.5). Diese dienten insbesondere dazu, die Befragten mit den Attributen und Konstrukten, deren Ausprägungen und der eingesetzten Skala vertraut zu machen. In einem Rating-Experiment wurden dem Befragten variiierende Ausprägungen der Attribute eines Konstrukts vorgelegt. Das durch die Kombination der Attributsausprägungen charakterisierte Konstrukt galt es, auf einer von „- - - -“ bis „+ + + +“ reichenden Skala zu bewerten. Um die Befragten

mit allen Ausprägungen, d. h. auch mit solchen, die möglicherweise nicht in den Rating-Experimenten des Befragten verwendet wurden, vertraut zu machen, waren jeweils alle Ausprägungen eines Konstrukts gleichzeitig sichtbar: Die in der entsprechenden Situation zu bewertenden Attributsausprägungen waren durch eine schwarze Schrift und hellblauem Hintergrund hervorgehoben, die anderen Ausprägungen hingegen hellgrau und ohne farbigen Hintergrund dargestellt.

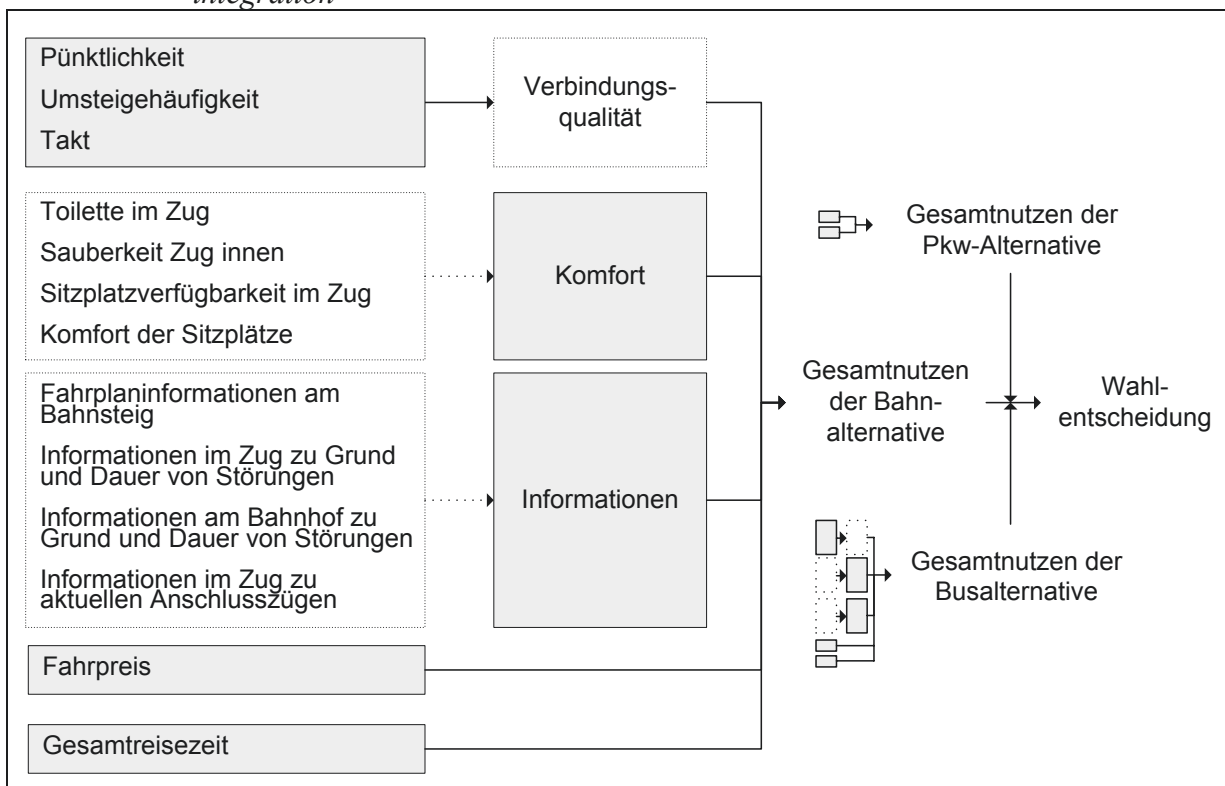
Die Kombination der Attributsausprägungen in den Rating-Experimenten erfolgte in Form eines zufälligen Designs (*random design*). Nach drei Bewertungen eines Konstrukts mit variierenden Attributsausprägungen wurde anschließend eine Bewertung der aktuellen Fahrt in Bezug auf dieses Konstrukt erfragt (Abb. D.6). Dieses Prozedere wurde für die anderen beiden Konstrukte wiederholt, sodass insgesamt neun hypothetische Konstruktbewertungen und drei Konstruktbewertungen für die aktuelle Fahrt erfolgten. Um systematische Reihenfolge- und Ermüdungseffekte zu vermeiden, wurde die Reihenfolge der drei Konstrukte in der Befragung zufällig variiert.

Nach den Rating-Experimenten folgten die Subexperimente der integrierten hierarchischen Informationsintegration (Abb. D.7 bis Abb. D.9). In diesen Wahlexperimenten konnten alle Befragten zwischen den beiden Alternativen Regionalbahn und Regionalbus wählen; Befragte, die die aktuelle Fahrt auch mit einem Pkw hätten durchführen können, konnten zudem eine Pkw-Alternative wählen. Die Wahl des Verkehrsmittels erfolgte mit Bezug zur aktuellen Fahrt, wobei sich die Befragten vorstellen sollten, dass für diese Fahrt nur die dargestellten Verkehrsmittel verfügbar waren.³⁸² Alle Verkehrsmittel wurden durch die Ausprägungen der Attribute *Gesamtreisezeit* und *Fahrpreis/Fahrtkosten* charakterisiert. Die Ausprägungen dieser beiden Attribute zur Beschreibung der beiden öffentlichen Verkehrsmittel wurden auf Basis der Eigenschaften der aktuellen Fahrt (mit der Regionalbahn) variiert; die Attribute zur Beschreibung der Pkw-Alternative wurden auf Basis der alternativ mit einem Pkw durchgeführten Fahrt variiert.

³⁸² Eine „no-option“ wurde nicht angeboten, da durch den Bezug zur aktuellen Fahrt angenommen wurde, dass eine grundsätzliche Entscheidung, diese Fahrt durchzuführen, bereits getroffen war und nur geringfügige und realistische Veränderungen der Attributsausprägungen vorgenommen wurden.

Zudem wurden die beiden Alternativen Regionalbahn und Regionalbus durch Ausprägungen der Attribute eines Konstrukts sowie hypothetische, zusammenfassende Konstruktwerte der beiden verbleibenden Konstrukte beschrieben. Die beiden Alternativen des öffentlichen Verkehrs wurden somit neben Zeit und Preis durch alle elf Qualitätsattribute mit Bezug zu *Verbindungsqualität*, *Komfort* und *Informationen* charakterisiert, entweder direkt als Attribut oder indirekt durch den zusammenfassenden Konstruktwert. Der Aufbau eines solchen Subexperiments ist in Abb. 7.2 dargestellt.

Abb. 7.2: *Aufbau des Subexperiments der integrierten hierarchischen Informationsintegration*



Quelle: *Eigene Darstellung.*

Die grau hinterlegten Attribute und Konstrukte werden im Subexperiment der Verbindungsqualität zur Beschreibung der Wahlalternativen verwendet.

Da jeweils immer eines der drei Konstrukte in die einzelnen Attribute aufgelöst wurde, ergaben sich folglich drei unterschiedliche Subexperimente. Innerhalb einer Befragung wurde immer das gleiche Konstrukt aufgelöst, sodass jeder Befragte nur mit einem der drei Subexperimente konfrontiert wurde.³⁸³ Die Auswahl der jeweiligen Ausprägungen in einer

³⁸³ Die Auswahl des Subexperiments erfolgte anhand der Reihenfolge der Rating-Experimente, d. h. wenn beispielsweise beim Rating-Experiment zuerst die Attribute

Wahlsituation erfolgte mit zufälligem Design (*random design*). Die Generierung der hierarchischen Wahlexperimente war als Endlosschleife programmiert, wobei ab fünf Wahlsituationen die Möglichkeit bestand, die Subexperimente der HII-I zu beenden und die Endlosschleife zu verlassen. Anschließend wurden zwei Brückenexperimente der HII (Abb. D.10) mit allen drei Konstrukten mit hypothetischen, zusammenfassenden Konstruktwerten durchgeführt.³⁸⁴

Die Fragen zum allgemeinen Verkehrsverhalten und zur Soziodemografie (Abb. D.11) bezogen sich auf die allgemeine Nutzungshäufigkeit der Bahn bzw. des Pkw (auch als Mitfahrer), den Besitz eines Pkw-Führerscheins, den Besitz eines Zeitfahrerscheins (z. B. Monatskarte, Semesterticket, Schwerbehindertenausweis), den Erhalt einer Fahrtkostenerstattung für die aktuelle Fahrt, das Geschlecht und das Alter des Befragten, die Anzahl der im Haushalt des Befragten lebenden Personen, ggf. die Anzahl der davon minderjährigen Personen sowie die Höhe des Bruttohaushaltseinkommens (unterteilt in grobe Kategorien). Abschließend erfolgte eine Danksagung (Abb. D.12). Der Interviewer hatte die Möglichkeit, eine Anmerkung zu hinterlassen (Abb. D.13).

Für die Befragung in der Nicht-Nutzer des regionalen SPNV (Datenerhebung in den Behörden) mussten Anpassungen des Erhebungsinstruments im Hinblick auf zwei Aspekte vorgenommen werden: Erstens war bei der Befragung in den Behörden eine Filterung zur Zielgruppenauswahl erforderlich, um sicherzustellen, dass die Befragten einen Bezug zum Regionalverkehr hatten. Zweitens war ein Bezug zu einer „aktuellen Fahrt“ nicht möglich.³⁸⁵ Der prinzipielle Ablauf dieses veränderten Erhebungsinstruments ist in Abb. 7.3 dargestellt, eine Übersicht der Fragebogenmasken befindet sich in Anhang D (Abb. E.1 bis Abb. E.10).

des Konstrukts der *Verbindungsqualität* bewertet wurden, wurde im Wahlexperiment dieses Konstrukt in die einzelnen Attribute aufgelöst.

³⁸⁴ Separat wurden auch Befragungen ohne Subexperimente der HII-I und ausschließlich mit Brückenexperimenten der HII durchgeführt. Wie bereits in Fußnote 381 erwähnt wird im Rahmen dieser Arbeit nicht weiter auf die Anwendung der HII eingegangen.

³⁸⁵ Ein Bezug zur Fahrt zur Behörde wäre grundsätzlich möglich gewesen. Davon wurde jedoch insbesondere aufgrund des stets gleichen (und von den Fahrtzwecken der *Bahnnutzer* abweichenden) Fahrtzwecks in Verbindung mit einem oftmals nicht bestehenden Bezug zum Regionalverkehr abgesehen.

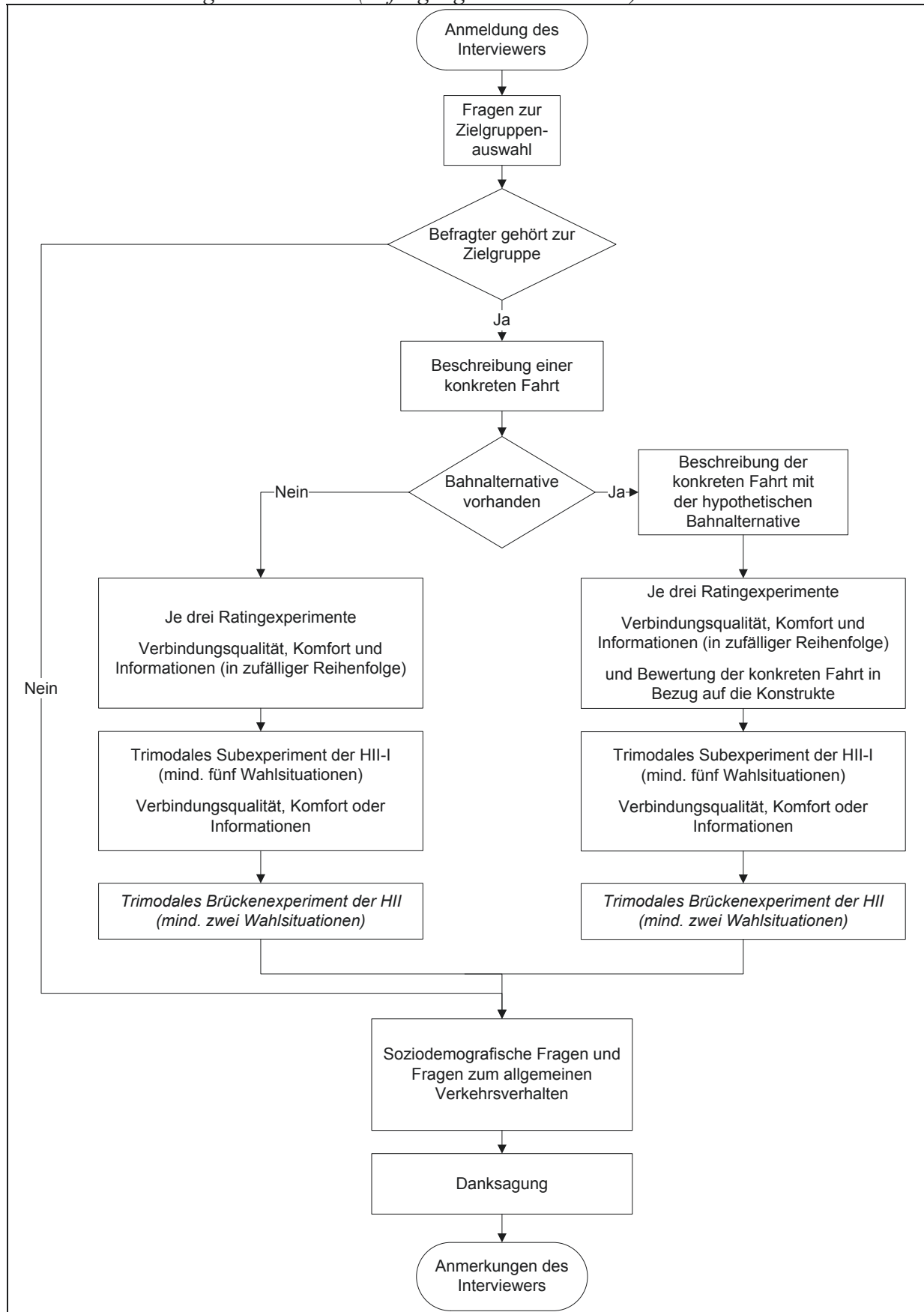
Für die Zielgruppenauswahl (Abb. E.2) wurden drei zu erfüllende Kriterien festgelegt: Die ersten beiden erfragten Kriterien stellten den Besitz eines Pkw-Führerscheins und die generelle Verfügbarkeit eines Pkw dar (d. h. der Pkw steht dem Befragten normalerweise *immer* oder *meistens* zur Verfügung). Das dritte erforderliche Kriterium war erfüllt, wenn der Befragte angab, dass er seine Arbeitsstätte mit dem Pkw (als Fahrer oder Mitfahrer) mit einer Fahrtzeit zwischen 20 und 90 Minuten erreichte und/oder regelmäßig sonstige Fahrten mit dem Pkw mit einer Fahrtzeit zwischen 20 und 90 Minuten durchführte. Sofern eines dieser drei Kriterien nicht erfüllt war, wurden keine Experimente durchgeführt und direkt die Fragen zum allgemeinen Verkehrsverhalten sowie zur Soziodemografie (Abb. E.8) gestellt.

Ansonsten folgte die Beschreibung einer konkreten, selbst auszuwählenden Fahrt (z. B. die Fahrt am Vortag zur Arbeit) (Abb. E.3). Hierbei wurden der Quelle-Ziel-bezogene Fahrtzweck, der Start- und Zielort, die Gesamtreisezeit (inklusive Weg zum Pkw, Parkplatzsuche, Stau, usw.), die geschätzten Fahrtkosten (jeweils für eine Richtung) sowie andere Kosten (z. B. Parkkosten) erfragt. Sofern bei der Zielgruppenauswahl das dritte Kriterium hinsichtlich der Fahrt zur Arbeitsstätte erfüllt war, wurden der Fahrtzweck sowie die Dauer der Fahrt direkt übernommen; Änderungen waren jedoch möglich (z. B. Beschreibung einer anderen Fahrt).

Die Fahrtkosten wurden bewusst nicht mit einem vorgegebenen Pauschalatz berechnet, da die erhobenen Werte die vom Befragten wahrgenommenen Kosten darstellen sollten, anhand derer auch in der Realität Entscheidungen getroffen werden. Dennoch wurde der Befragte, sofern erforderlich, bei der Errechnung bzw. Schätzung des Betrags durch den Interviewer unterstützt.

Des Weiteren wurde gefragt, ob die beschriebene Fahrt auch mit der Bahn hätte durchgeführt werden können. Falls diese Möglichkeit bestand, wurde nach dem Start- und dem Zielbahnhof, den verwendeten Verkehrsmitteln zum Erreichen und Verlassen der Bahnhöfe, der Umsteigehäufigkeit zwischen öffentlichen Verkehrsmitteln, sowie der Gesamtreisezeit und den Fahrtkosten gefragt.

Abb. 7.3: Aufbau des eingesetzten Erhebungsinstruments zur Befragung der Nicht-Nutzer des regionalen SPNV (Befragung in den Behörden)



Quelle: Eigene Darstellung.

Wie bei den Befragungen in den Zügen des Regionalverkehrs folgten anschließend die Rating-Experimente (Abb. E.4).³⁸⁶ Die Bewertungen der Konstrukte für die beschriebene Fahrt (Abb. E.5) wurden übersprungen, wenn die Fahrt nicht alternativ mit der Bahn hätte durchgeführt werden können.

In den Wahlexperimenten (Abb. E.6 bis Abb. E.7) standen immer alle drei Alternativen (Regionalbahn, Regionalbus und Pkw) zur Wahl, unabhängig davon, ob der Befragte die Bahn als mögliche Alternative für die beschriebene Fahrt angegeben hatte. Die Ausprägungen für die *Gesamtreisezeit* und *Fahrpreis/Fahrtkosten* wurden aus den Werten für die beschriebene Fahrt mit dem Pkw generiert. Die Attributsausprägungen der Pkw-Alternative schwankten dabei direkt um die für die Fahrt angegebenen Werte, wobei die *Fahrtkosten* auch die *sonstigen Kosten* enthielten. Die Ausprägungen der *Gesamtreisezeit* der Regionalbahn und des Regionalbusses schwankten um die mit dem Faktor 1,2 multiplizierte angegebene Gesamtreisezeit des Pkw (gerundet auf 5-Minuten-Intervalle) und die Ausprägungen des *Fahrpreises* um die mit dem Faktor 0,58 multiplizierten angegebenen Fahrtkosten des Pkw (ohne sonstige Kosten).³⁸⁷

Darüber hinaus wurden keine weiteren Anpassungen des Erhebungsinstruments vorgenommen.

Beide Erhebungsinstrumente wurden in einem Pretest in Form von Befragungen von Mitarbeitern und Studenten der Fachhochschule Gelsenkirchen auf die Verständlichkeit der Fragen hin überprüft. Die dabei erhobenen Daten wurden nicht in die Datenauswertung einbezogen.

Die Zielgruppe bei der Datenerhebung im Rahmen der Hauptstudie waren Nutzer und Nicht-Nutzer des regionalen SPNV. Die Durchführung der computergestützten persönlichen Interviews erfolgte überwiegend durch studentische Interviewer, welche auf ihren Einsatz in einer Schulung vorbereitet (vgl. Kap. 6.1) und bei ihrem ersten Einsatz begleitet wurden. Aus den gleichen Gründen wie bei der Vorstudie wurde die Datenerhe-

³⁸⁶ Ein Pretest zeigte, dass eine stärkere Betonung des fiktiven Charakters des Bahnangebots bei der Befragung von Pkw-Nutzern erforderlich war, welche daher erfolgte.

³⁸⁷ Diese beiden Faktoren orientierten sich an den Daten, die in den Zügen des Regionalverkehrs erhoben wurden.

bung als Ad-hoc-Stichprobe durchgeführt. Die Datenerhebung war in zwei Phasen unterteilt. In der ersten Phase fanden Befragungen von Nutzern des regionalen SPNV in den Zügen der Kooperationspartner statt. Die Datenerhebung erfolgte auf den Linien RB 50 (Münster – Dortmund), RB 66 (Münster – Osnabrück), RB 67 (Münster – Bielefeld) und RB 75 (Bielefeld – Osnabrück) im Zeitraum 12.05.2010 – 15.10.2010. Gegenstand der zweiten Phase waren Befragungen von Nicht-Nutzern des regionalen SPNV. In Analogie zur Vorstudie erfolgte die Befragung von Wartenden in Behörden, und zwar in den Kfz-Zulassungsstellen in Münster, Bielefeld und Osnabrück sowie im Bürgeramt der Stadt Münster im Zeitraum 29.12.2010 – 14.07.2011.³⁸⁸

Insgesamt wurden 3067 gültige Interviews mit Experimenten der integrierten hierarchischen Informationsintegration durchgeführt, davon 1609 in Zügen des Regionalverkehrs und 1458 in Behörden.

Bei der Befragung in den Zügen wurde während der Datenerhebung keine Filterung hinsichtlich der Gesamtreisezeit vorgenommen, sodass zunächst auch Daten von Fahrten mit einer Gesamtreisezeit von bis zu 10 Stunden erhoben wurden. Eine zeitliche Abgrenzung dergestalt, dass nur Daten von Fahrten mit einer Gesamtreisezeit von mehr als 15 Minuten und maximal 90 Minuten in die Auswertung einbezogen wurden, reduzierte die Anzahl der in den Zügen des Regionalverkehrs erhobenen Datensätze von 1609 auf 1192.³⁸⁹ Zudem wurden nur Daten von Bahnfahrten mit einem Fahrpreis von maximal 15 € sowie von (alternativen) Fahrten mit den Pkw mit Fahrtkosten von maximal 30 € einbezogen. Diese Filterung führte zu Datensätzen von 1152 Befragten in den Zügen, im Folgenden bezeichnet als *Bahnnutzer*, und 1432 Befragten in den Behörden, im Folgenden bezeichnet als *Pkw-Nutzer*. Die Gruppe der Bahnnutzer wurde zudem hinsichtlich der Pkw-Verfügbarkeit für die aktuelle Fahrt unter-

³⁸⁸ Zudem wurde der Versuch unternommen, Befragung mit Wartenden im Zentrum für Zahn-, Mund- und Kieferheilkunde des Universitätsklinikums Münster durchzuführen. Aufgrund der generell geringen Bereitschaft zur Teilnahme an der Befragung gepaart mit kurzen Wartezeiten der Wartenden, die vielfach zu einem Abbruch der Befragung führte, und einem relativ hohen Durchschnittsalter der Befragten wurde die Datenerhebung an diesem und einem vergleichbaren Ort nicht weiter verfolgt.

³⁸⁹ In 39 Befragungen wurde eine Gesamtreisezeit von 15 Minuten oder weniger und in 379 Befragungen eine Gesamtreisezeit von mehr als 90 Minuten angegeben.

schieden: Als *nicht wahlfrei* werden 661 Bahnnutzer ohne verfügbaren Pkw und als *wahlfrei* 491 Bahnnutzer mit verfügbarem Pkw bezeichnet.

7.2 Beschreibung der Stichprobe

Die Auswertung der Fragen zum allgemeinen Verkehrsverhalten sowie der soziodemografischen Charakteristika erfolgte differenziert nach Nutzergruppen. Eine Übersicht mit absoluten und adjustierten relativen Häufigkeiten ist in Tab. 7.2 dargestellt. Bei der Beschreibung der Stichprobe wird nicht nur eine Unterscheidung zwischen *nicht wahlfreien* und *wahlfreien Bahnnutzern*, sondern auch zwischen „*nicht wahlfreien*“ und „*wahlfreien*“ *Pkw-Nutzern* vorgenommen. Die Unterscheidung der Teilgruppen der *Pkw-Nutzer* ist für die spätere Analyse erforderlich und wird bereits bei der Beschreibung der Stichprobe berücksichtigt; die Zuordnung der *Pkw-Nutzer* zu den Teilgruppen wird jedoch erst in Kap. 7.5.1 erläutert.³⁹⁰

Die *nicht wahlfreien Bahnnutzer* waren im Durchschnitt 33,2 Jahre alt (SD: 13,8 J., Median: 29 J.), wobei 46,9 % der Alterskategorie *18 – 30 Jahre* angehörten. Die *wahlfreien Bahnnutzer*, „*nicht wahlfreien*“ *Pkw-Nutzer* und „*wahlfreien*“ *Pkw-Nutzer* gehörten mit 43,1 %, 46,9 % bzw. 45,2 % überwiegend der Alterskategorie *31 – 50 Jahre* an und waren durchschnittlich 37,7 Jahre (SD: 13,2 J., Median: 37 J.), 40,8 Jahre (SD: 13,6 J., Median: 40 J.) bzw. 37,1 Jahre (SD: 12,7, Median: 35 J.) alt.

Mit 60,9 %, 52,3 %, 55,2 % bzw. 56,5 % lebte in allen vier Gruppen die Mehrheit der Befragten in Ein- oder Zweipersonenhaushalten. Ungefähr die Hälfte (51,5 %) der *nicht wahlfreien Bahnnutzer* war weiblich. Von den *wahlfreien Bahnnutzern*, den „*nicht wahlfreien*“ *Pkw-Nutzern* und den „*wahlfreien*“ *Pkw-Nutzern* waren hingegen 61,1 %, 63,7 % bzw. 59,5 % männlich.

³⁹⁰ Diese Vorgehensweise wurde gewählt, um nicht zu einem späteren Zeitpunkt für eine Teilgruppe eine Beschreibung der Stichprobe in Bezug auf die Fragen zum allgemeinen Verkehrsverhalten sowie der soziodemografischen Charakteristika vornehmen zu müssen.

Tab. 7.2: Absolute und relative adjustierte Häufigkeiten der Charakteristika zur Soziodemografie und zum allgemeinen Verkehrsverhalten differenziert nach Nutzergruppen

| | Nicht wahlfreie Bahnnutzer (N=661) | | Wahlfreie Bahnnutzer (N=491) | | „Nicht wahlfreie“ Pkw-Nutzer (N=682) | | „Wahlfreie“ Pkw-Nutzer (N=750) | |
|--|--|---------------------|------------------------------------|---------------------|--|---------------------|--------------------------------------|---------------------|
| | Abs. | Rel. adj. (in %) | Abs. | Rel. adj. (in %) | Abs. | Rel. adj. (in %) | Abs. | Rel. adj. (in %) |
| Geschlecht | | | | | | | | |
| Männlich | 319 | 48,48 | 298 | 61,07 | 425 | 63,72 | 436 | 59,48 |
| Weiblich | 339 | 51,52 | 190 | 38,93 | 242 | 36,28 | 297 | 40,52 |
| Alter (Jahren) | | | | | | | | |
| ≤ 17 | 40 | 6,14 | 7 | 1,44 | 0 | 0,00 | 0 | 0,00 |
| 18 - 30 | 305 | 46,85 | 170 | 35,05 | 184 | 28,48 | 269 | 37,73 |
| 31 - 50 | 222 | 34,10 | 209 | 43,09 | 303 | 46,90 | 322 | 45,16 |
| ≥ 51 | 84 | 12,90 | 99 | 20,41 | 159 | 24,61 | 122 | 17,11 |
| Pkw-Führerscheinbesitz | | | | | | | | |
| Ja | 493 | 74,92 | 473 | 96,93 | 682 | 100,00 | 750 | 100,00 |
| Nein | 165 | 25,08 | 15 | 3,07 | | | | |
| Allg. Pkw-Verfügbarkeit | | | | | | | | |
| Immer | 79 | 12,02 | 306 | 62,70 | 651 | 95,45 | 691 | 92,13 |
| Meistens | 106 | 16,13 | 103 | 21,11 | 31 | 4,55 | 59 | 7,87 |
| Manchmal | 126 | 19,18 | 35 | 7,17 | | | | |
| Selten | 123 | 18,72 | 29 | 5,94 | | | | |
| Nie | 223 | 33,94 | 15 | 3,07 | | | | |
| Zeitfahrerscheinbesitz | | | | | | | | |
| Ja | 491 | 74,62 | 382 | 78,12 | 71 | 10,77 | 178 | 24,48 |
| Nein | 167 | 25,38 | 107 | 21,88 | 588 | 89,23 | 549 | 75,52 |
| Allg. Nutzungshäufigkeit Bahn | | | | | | | | |
| Mehrere Male in der | | | | | | | | |
| Woche | 463 | 70,36 | 380 | 77,71 | 10 | 1,51 | 37 | 5,08 |
| Mehrere Male im Monat | 127 | 19,30 | 53 | 10,84 | 42 | 6,33 | 96 | 13,19 |
| Selten | 62 | 9,42 | 53 | 10,84 | 424 | 63,86 | 480 | 65,93 |
| Nie | 6 | 0,91 | 3 | 0,61 | 188 | 28,31 | 115 | 15,80 |
| Allg. Nutzungshäufigkeit Pkw | | | | | | | | |
| Mehrere Male in der | | | | | | | | |
| Woche | 196 | 29,88 | 263 | 53,78 | 635 | 96,36 | 692 | 95,05 |
| Mehrere Male im Monat | 165 | 25,15 | 92 | 18,81 | 20 | 3,03 | 29 | 3,98 |
| Selten | 212 | 32,32 | 113 | 23,11 | 4 | 0,61 | 7 | 0,96 |
| Nie | 83 | 12,65 | 21 | 4,29 | 0 | 0,00 | 0 | 0,00 |
| Anzahl der Personen im Haushalt | | | | | | | | |
| 1 | 169 | 26,20 | 81 | 16,67 | 129 | 19,88 | 142 | 19,92 |
| 2 | 224 | 34,73 | 173 | 35,60 | 229 | 35,29 | 261 | 36,61 |
| 3 | 106 | 16,43 | 117 | 24,07 | 132 | 20,34 | 129 | 18,09 |
| 4 | 86 | 13,33 | 80 | 16,46 | 107 | 16,49 | 112 | 15,71 |
| 5 und mehr | 60 | 9,30 | 35 | 7,20 | 52 | 8,01 | 69 | 9,68 |

Quelle: Eigene Berechnungen.

Mit 74,9 %, 96,9 %, 100,0 % bzw. 100 % besaß in allen vier Gruppen eine deutliche Mehrheit der Befragten einen Pkw-Führerschein.³⁹¹ Über einen Zeitfahrchein verfügten 74,6 % der *nicht wahlfreien Bahnnutzer* und 78,1 % der *wahlfreien Bahnnutzer*, jedoch nur 10,8 % der „*nicht wahlfreien*“ *Pkw-Nutzer* und 24,5 % der „*wahlfreien*“ *Pkw-Nutzer*.

Ein derartiger Unterschied zwischen den Nutzergruppen ist auch bei der allgemeinen Nutzungshäufigkeit der Bahn zu erkennen: 70,4 % der *nicht wahlfreien* und 77,7 % der *wahlfreien Bahnnutzer* nutzten die Bahn *mehrmals in der Woche*, aber nur 1,5 % der „*nicht wahlfreien*“ und 5,1 % der „*wahlfreien*“ *Pkw-Nutzer*. Jedoch nutzten 92,2 % bzw. 81,7 % der beiden Gruppen der *Pkw-Nutzer* die Bahn *selten* oder *nie*. *Mehrmals in der Woche* nutzten den Pkw (auch als Mitfahrer) 29,9 % der *nicht wahlfreien Bahnnutzer*, 53,8 % der *wahlfreien Bahnnutzer*, 92,2 % der „*nicht wahlfreien*“ *Pkw-Nutzer* und 81,7 % der „*wahlfreien*“ *Pkw-Nutzer*. Hinsichtlich der allgemeinen Pkw-Verfügbarkeit gaben 33,9 % der *nicht wahlfreien Bahnnutzer* an, dass ihnen *nie* ein Pkw zur Verfügung stand und 62,7 % der *wahlfreien Bahnnutzer*, 95,5 % der „*nicht wahlfreien*“ *Pkw-Nutzer* und 92,1 % der „*wahlfreien*“ *Pkw-Nutzer*, dass ihnen *immer* ein Pkw zur Verfügung stand.³⁹² Aufgrund dieses Nutzungsverhaltens erschienen die Bezeichnungen *Bahnnutzer* und *Pkw-Nutzer* geeignet.

Eine Übersicht der Gesamtreisezeiten und Fahrtkosten der aktuellen³⁹³ bzw. alternativen Fahrten ist differenziert nach Nutzergruppen in Tab. 7.3 dargestellt. Die beiden Teilgruppen der *Bahnnutzer* einerseits und die beiden Teilgruppen der *Pkw-Nutzer* andererseits weisen jeweils annähernd gleiche durchschnittliche Gesamtreisezeiten und Fahrtkosten auf.

³⁹¹ Hierbei ist anzumerken, dass die Gruppe der *wahlfreien Bahnnutzer* auch Personen enthält, die die aktuelle Fahrt alternativ als Mitfahrer in einem Pkw hätten durchführen können. Bei der Gruppe der *Pkw-Nutzer* war der Besitz eines Pkw-Führerscheins hingegen eines der Kriterien der Zielgruppenauswahl.

³⁹² Bei der Befragung der *Pkw-Nutzer* dienten die Kategorien *immer* oder *meistens* der allgemeinen Pkw-Verfügbarkeit der Zielgruppenauswahl.

³⁹³ Aus Gründen der besseren Lesbarkeit wird die beschriebene Pkw-Fahrt der *Pkw-Nutzer* im Folgenden als aktuelle Fahrt bezeichnet.

Tab. 7.3: Gesamtreisezeiten und Fahrtkosten der aktuellen bzw. alternativen Fahrten differenziert nach Nutzergruppen

| | | | Nicht wahlfreie Bahnnutzer | Wahlfreie Bahnnutzer | „Nicht wahlfreie“ Pkw-Nutzer | „Wahlfreie“ Pkw-Nutzer |
|--|-------------------------------|----------|----------------------------------|-------------------------|------------------------------------|---------------------------|
| Gesamtreisezeit der aktuellen bzw. alternativen Fahrt (in min) | mit der Bahn bzw. mit dem Pkw | N | 661 | 491 | 201 | 258 |
| | | Mittelw. | 54,52 | 53,71 | 61,38 | 61,95 |
| | | SD | 20,58 | 20,31 | 39,06 | 36,85 |
| | | Median | 55,00 | 50,00 | 60,00 | 60,00 |
| Fahrtkosten der aktuellen bzw. alternativen Fahrt (in €) | mit der Bahn bzw. mit dem Pkw | N | | 491 | 682/201* | 750/258* |
| | | Mittelw. | | 44,28 | 34,51/40,78* | 33,57/40,66* |
| | | SD | | 18,01 | 17,44/19,31* | 16,65/18,84* |
| | | Median | | 45,00 | 30,00/40* | 30,00/35,00* |
| Fahrtkosten der aktuellen bzw. alternativen Fahrt (in €) | mit der Bahn bzw. mit dem Pkw | N | 661 | 491 | 156 | 228 |
| | | Mittelw. | 3,63 | 3,91 | 10,95 | 9,54 |
| | | SD | 3,22 | 3,14 | 9,93 | 9,39 |
| | | Median | 2,90 | 3,00 | 7,00 | 7,00 |
| Fahrtkosten der aktuellen bzw. alternativen Fahrt (in €) | mit der Bahn bzw. mit dem Pkw | N | | 491 | 682/157* | 750/229* |
| | | Mittelw. | | 8,75 | 6,57/8,36* | 5,96/7,81* |
| | | SD | | 5,14 | 5,77/6,46* | 4,98/5,68* |
| | | Median | | 8,00 | 5,00/6,00* | 5,00/6,00* |

Quelle: Eigene Berechnungen.

* Der vordere Wert bezieht sich auf alle „nicht wahlfreien“ bzw. „wahlfreien“ Pkw-Nutzer und der hintere Wert auf die „nicht wahlfreien“ bzw. „wahlfreien“ Pkw-Nutzer, die eine alternative Bahnfahrt beschrieben haben.

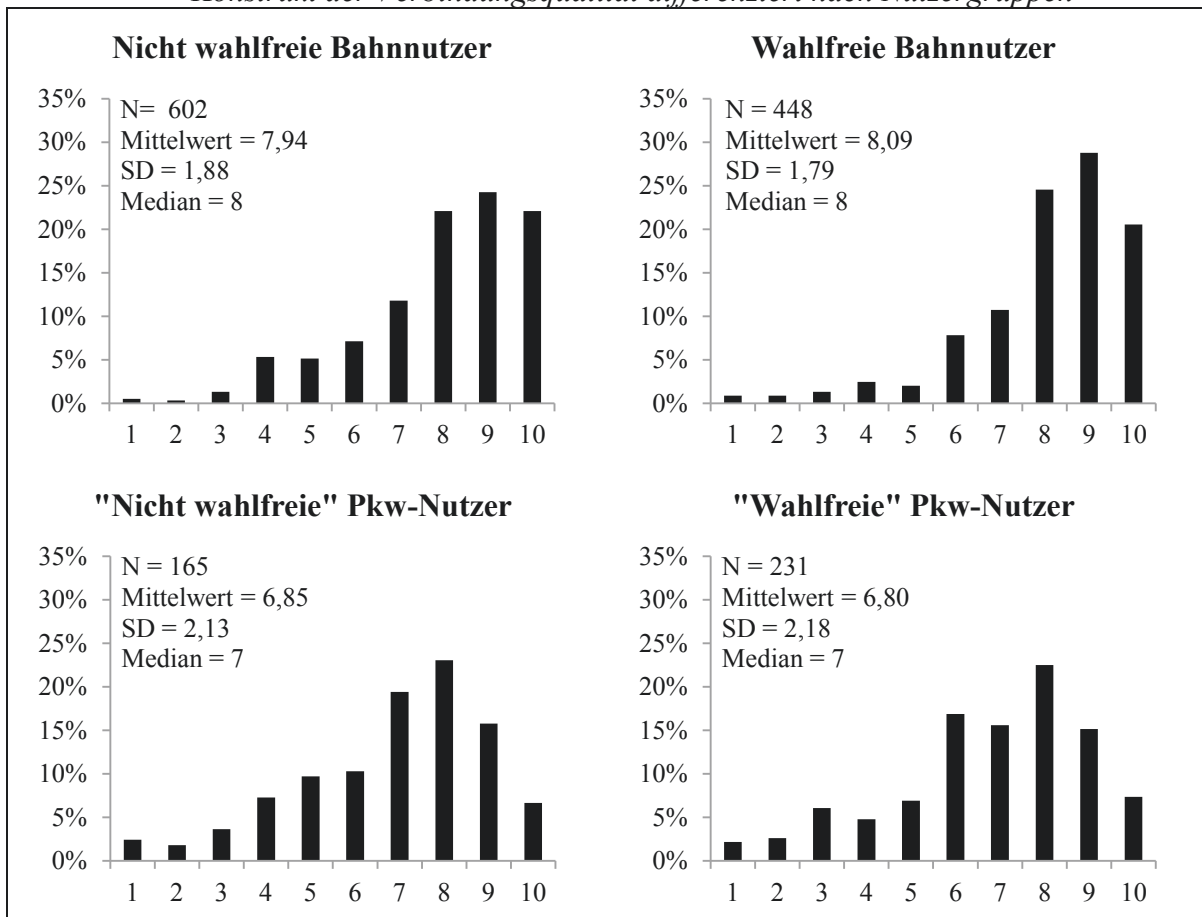
Die aktuelle Fahrt bezog sich bei allen vier Gruppen eher selten auf eine Wegekette.³⁹⁴ Bei 83,1 % der *nicht wahlfreien Bahnnutzer*, 92,7 % der *wahlfreien Bahnnutzer*, 94,1 % der „*nicht wahlfreien*“ *Pkw-Nutzer* und 96,8 % der „*wahlfreien*“ *Pkw-Nutzer* handelte es sich bei dem Start oder Ziel der Fahrt um die eigene Wohnung.

Als Quelle- oder Ziel-bezogenen Fahrtzweck der aktuellen Fahrt gaben 35,3 % der *nicht wahlfreien Bahnnutzer* den Beruf, 25,9 % die Ausbildung und 31,9 % die Freizeit an. Von den *wahlfreien Bahnnutzern* gaben 64,7 % den Beruf, 14,0 % die Ausbildung und ebenfalls 14,0 % die Freizeit als Quelle- oder Ziel-bezogenen Fahrtzweck der aktuellen Fahrt an. Von den Fahrten, die von den „*nicht wahlfreien*“ bzw. „*wahlfreien*“ *Pkw-Nutzern*

³⁹⁴ In den ersten Versionen des Erhebungsinstruments zur Befragung der *Bahnnutzer* wurde der Fahrtzweck allgemein erfragt, sodass es dem Befragten im Falle einer Wegekette überlassen war, den Hauptzweck der Fahrt zu bestimmen. In die folgende Beschreibung sind nur Daten einbezogen, bei denen der Fahrtzweck Quelle-Ziel-bezogen erhoben wurde.

beschrieben wurde, waren 57,9 % bzw. 58,3 % durch den Beruf, 1,3 % bzw. 2,7 % durch die Ausbildung und 29,8 % bzw. 30,1 % durch die Freizeit bedingt. Bei allen vier Gruppen betrafen nur (sehr) wenige der aktuellen Fahrten die Quelle-Ziel-bezogenen Fahrtzwecke Dienstreise, Einkauf und Sonstiges. Zusammenfassend lässt sich damit hervorheben, dass ein sehr hoher Anteil der aktuellen Fahrten der *wahlfreien Bahnnutzer* und der („nicht wahlfreien“ und „wahlfreien“) *Pkw-Nutzer*³⁹⁵ berufsbedingt ist, wohingegen die aktuellen Fahrten der *nicht wahlfreien Bahnnutzer* annähernd gleichermaßen durch den Beruf, die Ausbildung und die Freizeit bedingt sind. Die *Bahnnutzer* bewerteten die aktuelle Fahrt und die *Pkw-Nutzer* die alternative Fahrt mit der Bahn (sofern diese Alternative angegeben wurde) hinsichtlich der drei Konstrukte *Verbindungs-*

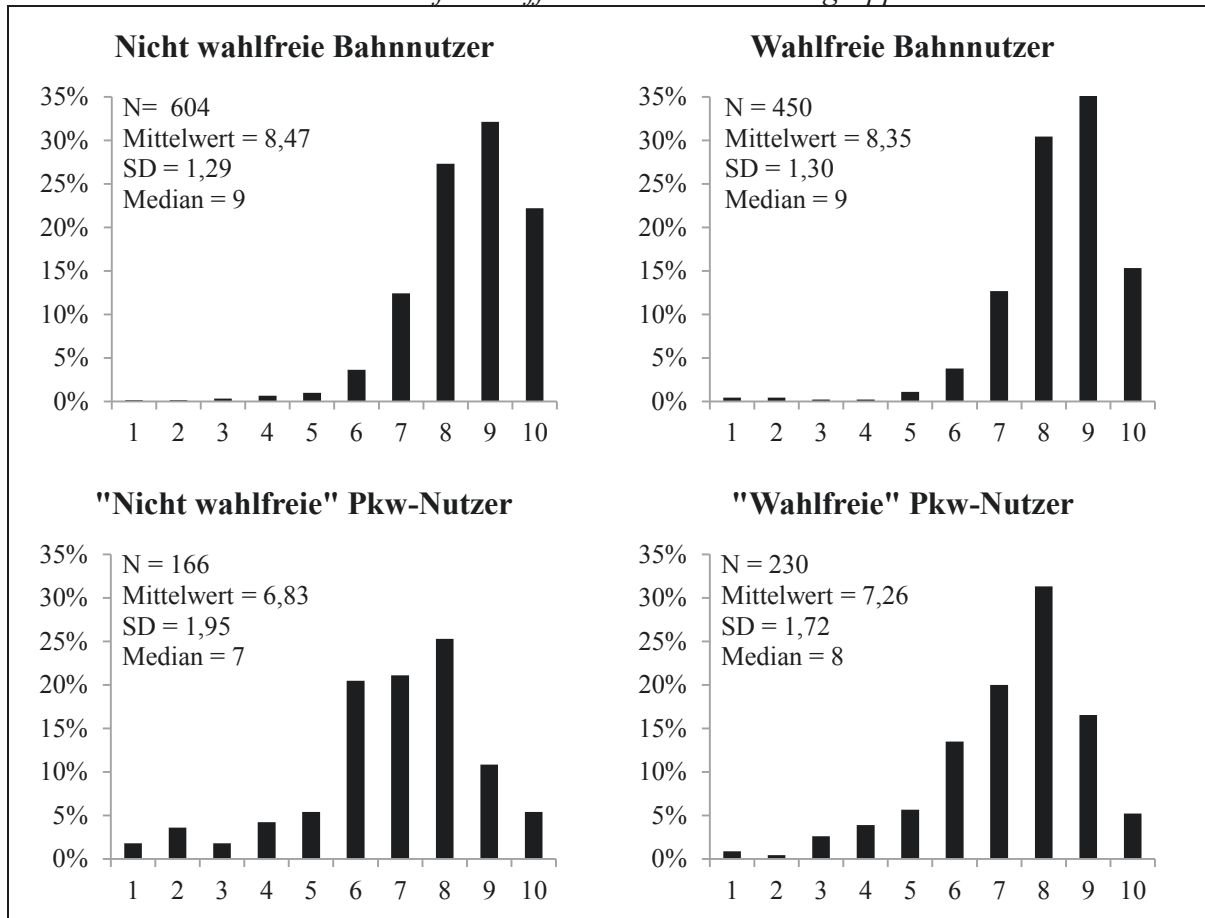
Abb. 7.4: Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt der Verbindungsqualität differenziert nach Nutzergruppen



Quelle: Eigene Erhebung.

³⁹⁵ Hinsichtlich der aktuellen Fahrten der *Pkw-Nutzer* sei anzumerken, dass der Quelle-Ziel-bezogene Fahrtzweck *von Wohnen nach Beruf* voreingestellt (aber dennoch veränderbar) war, sofern die Fahrt zur Arbeit mit dem Pkw mit der relevanten Gesamtreisezeit durchgeführt wurde.

Abb. 7.5: Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt des Komforts differenziert nach Nutzergruppen



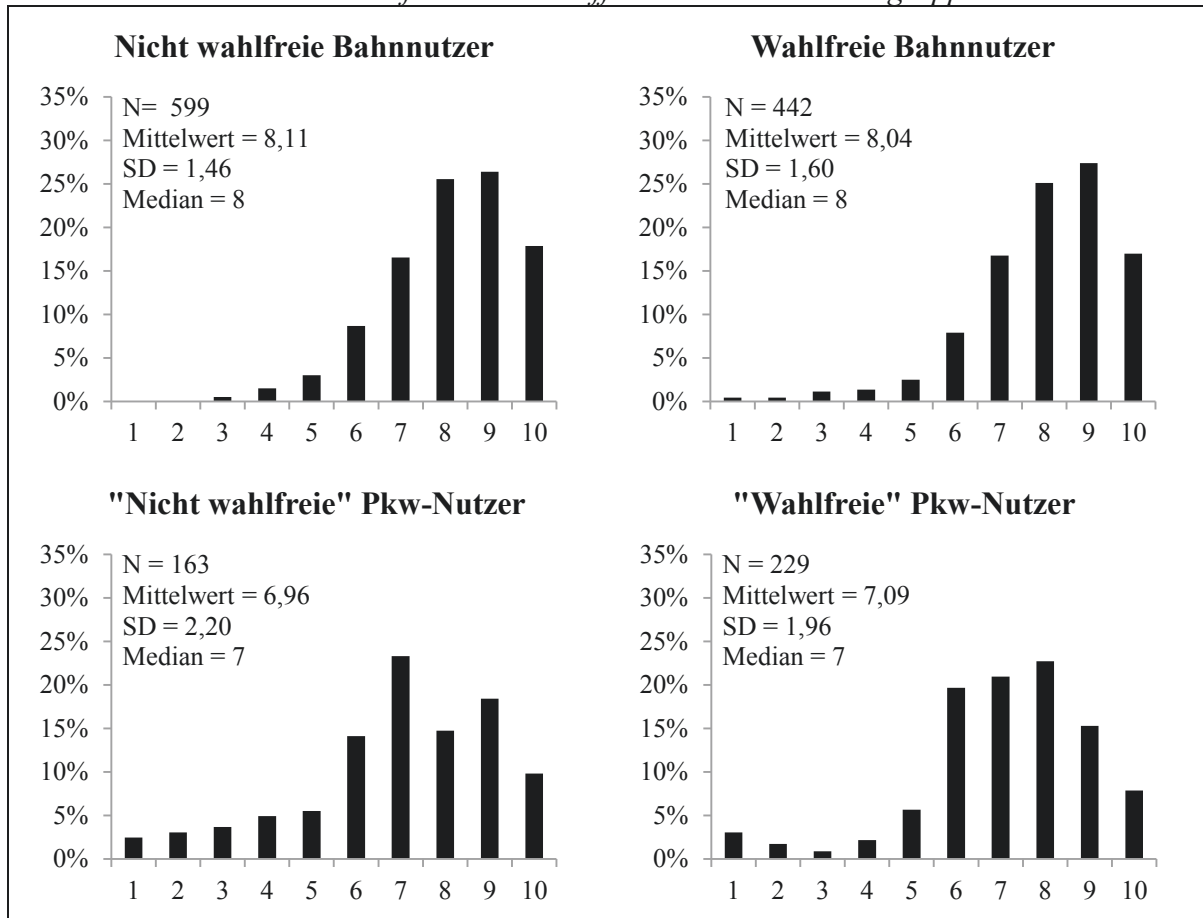
Quelle: Eigene Erhebung.

qualität, Komfort und Informationen. Die Bewertungen wurden numerisch kodiert, wobei der höchsten Bewertung „++++“ der Wert 10 und der niedrigsten Bewertung „-----“ der Wert 1 zugeordnet wurde. Die Verteilungen der Bewertungen der aktuellen bzw. alternativen Bahnfahrt in Bezug auf die drei Konstrukte ist differenziert nach Nutzergruppen in Abb. 7.4 bis Abb. 7.6 dargestellt. Im Allgemeinen wurde die aktuelle bzw. alternative Bahnfahrt in Bezug auf die drei Konstrukte von allen vier Nutzergruppen³⁹⁶ überwiegend positiv bis sehr positiv bewertet.

Anhand von Mediantests wurde überprüft, ob sich die Nutzergruppen hinsichtlich der Bewertungen der aktuellen bzw. alternativen Bahnfahrt unterscheiden. Die Testergebnisse zeigten, dass sich die beiden Gruppen der *Bahnnutzer* in Bezug auf die Bewertung des Konstrukts der *Verbindungsqualität* sowie in Bezug auf das Konstrukt der *Informationen*

³⁹⁶ Pkw-Nutzer ohne verfügbare Bahnalternative haben diese Bewertung nicht vorgenommen.

Abb. 7.6: Bewertung der aktuellen bzw. alternativen Bahnfahrt in Bezug auf das Konstrukt der Informationen differenziert nach Nutzergruppen



Quelle: Eigene Erhebung.

nicht unterscheiden, wohl aber in Bezug auf das Konstrukt des *Komforts*. Auch die beiden Gruppen der *Pkw-Nutzer* unterscheiden sich in Bezug auf die Bewertung des Konstrukts der *Verbindungsqualität* sowie in Bezug auf das Konstrukt der *Informationen* nicht, wohl aber in Bezug auf das Konstrukt des *Komforts*. In Bezug auf das Konstrukt des *Komforts* vergaben die *nicht wahlfreien Bahnnutzer* bzw. die „*wahlfreien*“ *Pkw-Nutzer* höhere Bewertungen. Zwischen den Gruppen der *Bahnnutzer* und der *Pkw-Nutzer* bestehen in Bezug auf alle drei Konstrukte Unterschiede, wobei die Gruppen der *Bahnnutzer* höhere Bewertungen vergaben.

7.3 Analyse der Rating-Experimente

Die Auswertung der Rating-Experimente erfolgte differenziert nach den vier Nutzergruppen anhand einer multiplen linearen Regressionsanalyse.³⁹⁷ Dabei stellten die Attribute die unabhängigen Variablen und die abgegebene Bewertung die abhängige Variable dar.

Zunächst war es erforderlich, eine Kodierung der qualitativen Attribute vorzunehmen. Statt der häufig eingesetzten Dummykodierung wurde hier die Effektkodierung verwendet. Die Wahl dieser Kodierungsvariante erfolgte im Sinne einer einheitlichen Kodierung bei allen Rating- und Wahlexperimenten; bei der Analyse der Wahlexperimente bietet die Effektkodierung im Gegensatz zur Dummykodierung den Vorteil, dass keine Vermischung der Effekte der Referenzkategorien der qualitativen Attribute mit der sogenannten alternativenspezifischen Konstante erfolgt.³⁹⁸

Bei der Effektkodierung eines Attributs mit L Ausprägungen werden $L-1$ Indikatorvariablen verwendet, wobei

- allen Indikatorvariablen der Wert -1 zugeordnet wird, wenn die Ausprägung der Referenzkategorie vorliegt,
- der l -ten Indikatorvariablen der Wert 1 zugeordnet wird, wenn die l -te Ausprägung des Attributs (mit $l = 1, \dots, L-1$) vorliegt und
- den Indikatorvariablen bei allen anderen Ausprägungen der Wert 0 zugewiesen wird.

Im Gegensatz zur Dummykodierung kann bei der Effektkodierung für die Referenzkategorie ein eigener Wert ermittelt werden, welcher der negativen Summe der Indikatorvariablen der anderen Ausprägungen entspricht.

Die Effektkodierung ist exemplarisch für das Attribut *Pünktlichkeit* mit der Referenzkategorie *pünktlich bzw. max. 3 Minuten Verspätung* in Tab. 7.4 dargestellt. Ebenso wurden die anderen qualitativen Attribute

³⁹⁷ Ebenso wie bei der Beschreibung der Stichprobe wurde nicht nur zwischen *nicht wahlfreien* und *wahlfreien Bahnutzern*, sondern auch zwischen „*nicht wahlfreien*“ und „*wahlfreien*“ *Pkw-Nutzern* unterschieden.

³⁹⁸ Vgl. *Bech/Gyrd-Hansen* (2005), S. 1080 und *Adamowicz/Louviere/Williams* (1994), S. 280–281.

Tab. 7.4: Effektkodierung des Attributs Pünktlichkeit

| Attributsausprägungen | Indikatorvariablen | |
|---|-----------------------------|-----------------------------|
| | X _{Pünktlichkeit2} | X _{Pünktlichkeit3} |
| 1. pünktlich bzw. max. 3 Minuten Verspätung | -1 | -1 |
| 2. 3 bis 10 Minuten Verspätung | 1 | 0 |
| 3. 10 bis 20 Minuten Verspätung | 0 | 1 |

Quelle: Eigene Darstellung.

kodiert, wobei als Referenzkategorie jeweils die oberste/beste³⁹⁹ Ausprägung gewählt wurde.

Abweichend davon wurden bei den drei Informationsattributen mit den Ausprägungen *keine Informationen*, *Anzeige*, *Durchsage* und *Anzeige und Durchsage* getrennt nach den Arten der Informationsvermittlung die Effekte *Anzeige* und *Durchsage* sowie die Interaktion zwischen diesen beiden Effekten kodiert.⁴⁰⁰ Den Konstruktbewertungen wurden numerische Werte von 1 („- - - -“) bis 10 („+ + + +“) zugeordnet und die Annahme äquidistanter Abstände zwischen den einzelnen Ausprägungen wurde getroffen (vgl. Kap. 3.2).

In Tab. 7.5 bis Tab. 7.7 sind die Ergebnisse der Regressionsanalysen differenziert nach den drei Konstrukten und den vier Nutzergruppen dargestellt. Ausgehend von den geschätzten Parametern wurden die Parameter der jeweiligen Referenzkategorien berechnet. Zudem ist in den Tabellen das Bestimmtheitsmaß R^2 angegeben, welches den Anteil der erklärten Streuung an der Gesamtstreuung beschreibt.

Bei der Regressionskonstante handelt es sich um den ungewichteten Gesamtmittelwert der Konstruktbewertungen. Die jeweiligen Regressionsparameter wurden als Abweichung von dieser Konstante interpretiert: Attributsausprägungen mit einem positiven Parameter sind mit überdurchschnittlich hohen Bewertungen verbunden, Attributsausprägungen mit einem negativen Parameter mit unterdurchschnittlich niedrigen Bewertungen; Bewertungen von Attributsausprägungen mit nicht signifi-

³⁹⁹ Diese Ausprägung ist in Tab. 7.1 mit „1.“ gekennzeichnet.

⁴⁰⁰ Ein vorhandener Effekt bzw. eine vorhandene Interaktion wurde dabei mit 1 und ein nicht vorhandener Effekt bzw. eine nicht vorhandene Interaktion mit -1 kodiert. Diese nach den Arten der Informationsvermittlung getrennte Effektkodierung führt zu identischen Ergebnissen wie eine „normale“ Effektkodierung.

Tab. 7.5: Regressionsmodell der Verbindungsqualität differenziert nach Nutzergruppen

| | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | | „Nicht wahlfreie“ Pkw-Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|---|-------------------------------|-----------|-------------------------|-----------|------------------------------------|-----------|---------------------------|-----------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Pünktlichkeit | | | | | | | | |
| 1. <i>pünktlich</i> bzw. <i>max.</i> <i>3 Minuten</i> <i>Verspätung</i> | 1,191 | | 1,012 | | 1,109 | | 1,185 | |
| 2. 3 bis 10 Minuten Verspätung | -0,004 | (-0,07) | 0,044 | (0,56) | -0,041 | (-0,54) | -0,074 | (-1,14) |
| 3. 10 bis 20 Minuten Verspätung | -1,186 | (-19,09) | -1,056 | (-13,36) | -1,069 | (-14,29) | -1,111 | (-16,80) |
| Umsteigehäufigkeit | | | | | | | | |
| 1. <i>0 x Umsteigen</i> | 0,836 | | 1,050 | | 0,886 | | 0,871 | |
| 2. 1 x Umsteigen mit Anschlussgarantie | 0,297 | (4,70) | 0,188 | (2,37) | 0,150 | (2,00) | 0,222 | (3,38) |
| 3. 1 x Umsteigen ohne Anschlussgarantie | -1,133 | (-18,07) | -1,238 | (-15,69) | -1,036 | (-13,73) | -1,093 | (-16,71) |
| Takt | | | | | | | | |
| 1. <i>alle 30 Minuten</i> | 1,577 | | 1,632 | | 0,816 | | 1,185 | |
| 2. <i>alle 60 Minuten</i> | 0,224 | (3,56) | 0,133 | (1,67) | 0,174 | (2,34) | 0,008 | (0,13) |
| 3. <i>alle 120 Minuten</i> | -1,802 | (-28,67) | -1,765 | (-22,51) | -0,989 | (-13,23) | -1,193 | (-18,07) |
| Konstante | 5,327 | (120,07) | 5,191 | (93,56) | 5,198 | (98,04) | 5,553 | (119,74) |
| <i>N</i> | 1965 | | 1453 | | 2000 | | 2232 | |
| <i>R</i> ² | 0,480 | | 0,449 | | 0,257 | | 0,350 | |
| adj. <i>R</i> ² | 0,479 | | 0,447 | | 0,255 | | 0,349 | |

Quelle: Eigene Berechnungen.

kantem Parameter unterscheiden sich nicht von den durchschnittlichen Bewertungen.

Sofern für keine der Ausprägungen eines Attributs ein signifikanter Regressionsparameter geschätzt wird, kann folglich geschlossen werden, dass dieses Attribut die Konstruktbewertung nicht einheitlich systematisch beeinflusst.

Die Auswertung der Ratingdaten der vier Gruppen zeigt, dass bei allen Attributen für mindestens eine Ausprägung ein signifikanter Regressionsparameter geschätzt wurde.⁴⁰¹ Daraus wurde gefolgert, dass die Befragten auf Veränderungen der Attributsausprägungen bei der Bewertung der Konstrukte reagiert haben. Ein grundsätzlicher Zusammenhang zwischen diesen Attributen und Konstrukten ist somit gegeben.

⁴⁰¹ Die geschätzten Parameter sind signifikant, wenn die *t*-Statistik den kritischen Wert $t_{krit} = 1,96$ (Signifikanzniveau von $\alpha = 0,05$) betragsmäßig überschreitet.

Tab. 7.6: Regressionsmodell des Komforts differenziert nach Nutzergruppen

| | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | | „Nicht wahlfreie“ Pkw-Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|--|-------------------------------|-----------|-------------------------|-----------|------------------------------------|-----------|---------------------------|-----------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Toilette im Zug/Bus | | | | | | | | |
| 1. sauber | 0,901 | | 0,920 | | 1,175 | | 1,027 | |
| 2. verschmutzt | -0,456 | (-7,49) | -0,270 | (-3,78) | -0,599 | (-8,51) | -0,538 | (-8,52) |
| 3. nicht vorhanden | -0,445 | (-7,31) | -0,651 | (-8,97) | -0,576 | (-8,34) | -0,490 | (-7,81) |
| Sauberkeit Zug/Bus innen | | | | | | | | |
| 1. sauber | 1,010 | | 1,047 | | 0,923 | | 0,992 | |
| 2. Boden klebrig, Papier liegt auf den Sitzen | -1,010 | (-23,35) | -1,047 | (-20,63) | -0,923 | (-18,83) | -0,992 | (-22,36) |
| Sitzplatzverfügbarkeit im Zug/Bus | | | | | | | | |
| 1. ganze Fahrt Sitzplatz | 1,606 | | 1,540 | | 1,166 | | 1,199 | |
| 2. halbe Fahrt Sitzplatz | -0,053 | (-0,87) | -0,227 | (-3,15) | -0,074 | (-1,06) | 0,001 | (0,01) |
| 3. kein freier Sitzplatz | -1,553 | (-25,14) | -1,313 | (-18,29) | -1,092 | (-15,63) | -1,200 | (-19,03) |
| Komfort der Sitzplätze | | | | | | | | |
| 1. bequem | 0,623 | | 0,679 | | 0,537 | | 0,583 | |
| 2. unbequem | -0,623 | (-14,38) | -0,679 | (-13,36) | -0,537 | (-10,96) | -0,583 | (-13,14) |
| Konstante | 4,558 | (105,35) | 4,563 | (89,88) | 4,276 | (87,26) | 4,688 | (105,70) |
| N | 1964 | | 1456 | | 1998 | | 2229 | |
| R ² | 0,492 | | 0,497 | | 0,356 | | 0,393 | |
| adj. R ² | 0,490 | | 0,495 | | 0,354 | | 0,391 | |

Quelle: Eigene Berechnungen.

Nicht für alle effektkodierten Indikatorvariablen der Attribute *Pünktlichkeit*, *Takt* und *Sitzplatzverfügbarkeit* wurden signifikante Parameter geschätzt. Der nicht signifikante Parameter bezieht sich dabei allerdings stets auf die mittlere von insgesamt drei Attributsausprägungen; für die untere Ausprägung wurde hingegen ein signifikanter Parameter geschätzt. Daraus wurde gefolgert, dass ein linearer Zusammenhang zwischen diesen Attributsausprägungen und den Konstruktbewertungen besteht.

Die Parameter der einzelnen Attributsausprägungen weisen plausible Vorzeichen und Rangfolgen auf, d. h. die jeweils oberste/beste Ausprägung führt zu einer höheren Bewertung als die unterste/schlechteste Ausprägung eines Attributs.

Die Rangfolge der zweiten und dritten Ausprägung des Attributs *Toilette im Zug* ist je nach Nutzergruppe unterschiedlich. Während die *wahlfreien Bahnnutzer* eine *verschmutzte* Toilette besser bewerteten als eine *nicht vorhandene* Toilette, bewerteten die *nicht wahlfreien* und *wahlfreien Pkw-*

Tab. 7.7: Regressionsmodell der Informationen differenziert nach Nutzergruppen

| | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | | „Nicht wahlfreie“ Pkw- Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|---|-------------------------------|-----------|-------------------------|-----------|-------------------------------------|-----------|---------------------------|-----------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Fahrplaninformationen am Bahnhof | | | | | | | | |
| 1. vorhanden | 0,933 | | 0,789 | | 0,945 | | 1,058 | |
| 2. nicht vorhanden | -0,933 | (-21,53) | -0,789 | (-15,57) | -0,945 | (-19,24) | -1,058 | (-23,97) |
| Informationen im Zug zu Grund und Dauer von Störungen | | | | | | | | |
| Durchsageneffekt | 0,645 | (14,90) | 0,592 | (11,70) | 0,447 | (9,11) | 0,449 | (10,18) |
| Anzeigeneffekt | 0,450 | (10,40) | 0,518 | (10,23) | 0,397 | (8,11) | 0,470 | (10,66) |
| Interaktionseffekt | -0,364 | (-8,39) | -0,323 | (-6,37) | -0,296 | (-6,03) | -0,368 | (-8,34) |
| Informationen am Bahnhof zu Grund und Dauer von Störungen | | | | | | | | |
| Durchsageneffekt | 0,672 | (15,53) | 0,670 | (13,18) | 0,440 | (8,97) | 0,490 | (11,12) |
| Anzeigeneffekt | 0,599 | (13,84) | 0,664 | (13,05) | 0,503 | (10,24) | 0,499 | (11,31) |
| Interaktionseffekt | -0,348 | (-8,03) | -0,422 | (-8,31) | -0,330 | (-6,72) | -0,339 | (-7,68) |
| Informationen im Zug zu aktuellen Anschlusszügen | | | | | | | | |
| Durchsageneffekt | 0,561 | (12,96) | 0,535 | (10,58) | 0,595 | (12,13) | 0,420 | (9,51) |
| Anzeigeneffekt | 0,460 | (10,62) | 0,424 | (8,37) | 0,548 | (11,17) | 0,568 | (12,86) |
| Interaktionseffekt | -0,337 | (-7,79) | -0,330 | (-6,51) | -0,315 | (-6,40) | -0,384 | (-8,69) |
| Konstante | 5,734 | (132,59) | 5,882 | (115,82) | 5,547 | (113,15) | 5,779 | (130,96) |
| <i>N</i> | 1951 | | 1458 | | 1996 | | 2233 | |
| <i>R</i> ² | 0,461 | | 0,445 | | 0,362 | | 0,423 | |
| adj. <i>R</i> ² | 0,458 | | 0,441 | | 0,359 | | 0,421 | |

Quelle: Eigene Berechnungen.

Nutzer sowie die nicht wahlfreien Bahnnutzer eine nicht vorhandene Toilette besser als eine verschmutzte Toilette.

Die Parameter, die für die Interaktionen zwischen *Anzeige* und *Durchsage* der jeweiligen Informationsattribute geschätzt wurden, weisen negative Vorzeichen auf. Dies bedeutet zum einen, dass bei dem gleichzeitigen Vorhandensein von *Anzeige* und *Durchsage* die einzelnen (positiven) Effekte durch die Interaktion abgeschwächt werden, sodass die Konstruktbewertung um einen Wert ansteigt, der kleiner ist als die Summe der Parameter der beiden einzelnen Effekte. Daraus lässt sich eine teilweise Redundanz der beiden Arten der Informationsvermittlung folgern. Zum anderen bedeutet die Interaktion, dass bei dem gleichzeitigen Nichtvorhandensein von *Anzeige* und *Durchsage* die einzelnen (negativen) Effekte verstärkt werden, sodass die Konstruktbewertung um einen Wert sinkt, der größer ist als die Summe der Parameter der beiden einzelnen Effekte.

Bei der Durchführung der Rating-Experimente waren jeweils alle Attributsausprägungen gleichzeitig sichtbar (Abb. D.3 bis Abb. D.5). Eine derartige Anordnung der Attributsausprägungen hätte grundsätzlich dazu führen können, dass der Befragte vereinfachende Strategien bei der Bewertung dergestalt verwendet, dass er weniger auf die Attributsausprägung an sich, sondern vielmehr auf deren Position auf dem Bildschirm (auf der linken Seite, in der Mitte bzw. auf der rechten Seite) achtet. Auf eine solche vereinfachende Strategie deuten durchgehend lineare Zusammenhänge zwischen den Attributsausprägungen und Konstruktbewertungen hin. Wie oben beschrieben wurden lineare Zusammenhänge allerdings nur vereinzelt festgestellt.

7.4 Aufbau der Analyse der hierarchischen Wahlexperimente

In den Wahlexperimenten der integrierten hierarchischen Informationsintegration wurde jeweils eines der drei Konstrukte *Verbindungsqualität*, *Komfort* und *Informationen* in seine einzelnen Attribute aufgelöst. Im Folgenden wird abhängig davon, welches Konstrukt aufgelöst wurde, zwischen den drei Subexperimenten unterschieden. Zudem erfolgt eine Differenzierung der Nutzergruppen nach Erhebungsort (in den Zügen des Regionalverkehrs bzw. in den Behörden) und der Anzahl der Wahlalternativen (Unterscheidung nach realer Pkw-Verfügbarkeit), sodass (zunächst) zwischen den *nicht-wahlfreien Bahnnutzern*, den *wahlfreien Bahnnutzern* und den *Pkw-Nutzern* unterschieden wird.

Für die einzelnen Nutzergruppen und Subexperimente werden zunächst unter Verwendung des multinomialen Logit-Modells mehrere getrennte Nutzenfunktionen geschätzt. Ausgehend von diesem einfachen Modell liegen der folgenden Analyse im Wesentlichen drei Fragestellungen zugrunde:

1. Besteht Gleichheit zwischen den Nutzenfunktionen der Nutzergruppen?
2. Wird die dem multinomialen Logit-Modell zugrunde liegende Annahme der Unabhängigkeit von irrelevanten Alternativen (IIA) nicht verletzt?
3. Besteht Prozessgleichheit zwischen den drei Subexperimenten?

Getrennt für die drei Subexperimente wird mit Testverfahren überprüft, ob die Nutzenfunktionen der Nutzergruppen gleich sind. Sofern keine Unterschiede bestehen erfolgt eine gemeinsame Analyse, andernfalls eine nach Nutzergruppen getrennte Analyse.

Das multinomiale Logit-Modell basiert auf der Annahme, dass die nicht beobachteten Effekte ε der Alternativen unabhängig voneinander und identisch verteilt sind und somit die Annahme der Unabhängigkeit von irrelevanten Alternativen besteht. Ist diese Annahme verletzt, so führt eine Analyse mit dem multinomialen Logit-Modell zu verzerrten Ergebnissen (vgl. Kap. 3.2). In diesem Fall stellt das Nested Logit-Modell eine mögliche Modellalternative dar, da dieses Modell auf einer partiellen Lockerung der IID-Annahme basiert, dergestalt, dass nur für Subsets von Alternativen eine voneinander unabhängige und identische Verteilung der nicht beobachteten Effekte ε getroffen wird. Somit können mehrere Alternativen gemeinsame nicht beobachtete Effekte ε teilen (vgl. Kap. 3.3).

Im Rahmen der zweiten Fragestellung wird daher die Annahme der Unabhängigkeit von irrelevanten Alternativen differenziert nach den drei Subexperimenten und ggf. zusätzlich differenziert nach den Nutzergruppen getestet. Sofern diese Annahme nicht verletzt ist, wird das MNL-Modell für die weitere Analyse beibehalten. Andernfalls wird das MNL-Modell verworfen und die weitere Analyse mit einem NL-Modell durchgeführt. Das Testverfahren der ersten Fragestellung (Gleichheit der Nutzenfunktionen der Nutzergruppen) wird in diesen Fall mit dem neuen Modell wiederholt.

Als Ergebnis der ersten beiden Fragestellungen steht somit fest, welches Modell verwendet wird und ob die Analyse für die Nutzergruppen gemeinsam oder getrennt durchgeführt wird. Aufbauend auf diesen Ergebnissen erfolgt das Testverfahren zur dritten Fragestellung. Hierbei wird analysiert, ob zwischen den drei Subexperimenten Prozessgleichheit besteht, d. h. ob die unterschiedlichen Subexperimente das gleiche Wahlverhalten abbilden. Prozessgleichheit ist gegeben, wenn die Nutzenfunktionen zwischen den drei Subexperimenten gleich sind. Dieser Zusammenhang lässt sich wie folgt begründen: In jedem der drei Subexperimente wurde genau ein Konstrukt in die Attribute zerlegt. Ändern sich die Nutzenfunktionen durch die Auflösung eines Konstrukts in seine Attribute

nicht, bedeutet dies, dass das Konstrukt die Attribute richtig abbildet. Somit handelt es sich bei der Überprüfung der Prozessgleichheit um die Überprüfung des hierarchischen Modells für diese konkrete Anwendung. Sofern Prozessgleichheit zwischen den drei Subexperimenten besteht, können im Ergebnis die Modelle aller drei Subexperimente zu einem gemeinsamen Modell vereint werden, d. h. auf Basis der Daten der drei Subexperimente kann ein geeignetes gemeinsames Modell geschätzt werden (vgl. Kap. 4.3.3).

7.5 Multinomiales Logit-Modell

Im Sinne einer einheitlichen Kodierung wurde eine Effektkodierung der qualitativen Attribute wie in Kap. 7.3 beschrieben vorgenommen. Zudem wurden die Ausprägungen der hypothetischen, zusammenfassenden Konstruktwerte „++++“, „++“ und „-“ mit 9, 7 bzw. 4 numerisch kodiert.⁴⁰² Die Gesamtreisezeiten und Fahrpreise/Fahrtkosten wiesen bereits numerische Werte auf, sodass für diese Attribute keine Kodierung vorgenommen wurde.

Allgemein werden mithilfe von alternativenspezifischen Konstanten Unterschiede zwischen den Wahlanteilen und somit im Nutzen, den die Alternativen stiften, berücksichtigt, die nicht durch die Attribute der Wahlalternativen erfasst werden. Bei J Alternativen ist eine Parameterschätzung für bis zu $J-1$ alternativenspezifische Konstanten möglich.⁴⁰³ Die Parameter dieser Konstanten sind im Vergleich zu einer Referenzalternative (d. h. der Alternative ohne Konstante) zu interpretieren. Grundsätzlich ist es beliebig, welchen Alternativen jeweils eine Konstante zugeordnet wird und welche Alternative als Referenzalternative dient. Im Rahmen dieser Arbeit war die Pkw-Alternative jedoch nicht in allen Experimenten verfügbar und somit als einheitliche Referenzalternative aller Experimente ungeeignet. Der Regionalbus diente als Referenzalternative, sodass der Regionalbahn und dem Pkw jeweils eine alternativenspezifische Konstante in Form einer Indikatorvariablen zugeordnet wurde.

⁴⁰² Eine Effektkodierung der Konstrukte führte im Rahmen der Analyse zu keiner signifikanten Verbesserungen der Modelle.

⁴⁰³ Vgl. *Louviere et al.* (2000), S. 49 und *Hensher/Rose/Greene* (2005), S. 312–315.

7.5.1 Subexperiment der Verbindungsqualität

Für das Subexperiment der *Verbindungsqualität* wurden differenziert nach den Nutzergruppen multinomiale Logit-Modelle geschätzt.⁴⁰⁴ Die den Modellen zugrunde liegenden Nutzenfunktionen der Bahn-, Bus- und Pkw-Alternative lauten

$$V(\text{Bahn}) =$$

$$\begin{bmatrix} \beta_{\text{Pünktlichkeit2}} \\ \beta_{\text{Pünktlichkeit3}} \\ \beta_{\text{Takt2}} \\ \beta_{\text{Takt3}} \\ \beta_{\text{Umsteigen2}} \\ \beta_{\text{Umsteigen3}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Pünktlichkeit2}} \\ x_{\text{Pünktlichkeit3}} \\ x_{\text{Takt2}} \\ x_{\text{Takt3}} \\ x_{\text{Umsteigen2}} \\ x_{\text{Umsteigen3}} \end{bmatrix} + \begin{bmatrix} \gamma_{\text{Bahn}} \\ \gamma_{\text{ZeitÖV}} \\ \gamma_{\text{PreisÖV}} \\ \gamma_{\text{Komfort}} \\ \gamma_{\text{Inform.}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Bahn}} \\ x_{\text{Zeit}} \\ x_{\text{Preis}} \\ x_{\text{Komfort}} \\ x_{\text{Inform.}} \end{bmatrix}, \quad (7.1)$$

$$V(\text{Bus}) = \begin{bmatrix} \beta_{\text{Pünktlichkeit2}} \\ \beta_{\text{Pünktlichkeit3}} \\ \beta_{\text{Takt2}} \\ \beta_{\text{Takt3}} \\ \beta_{\text{Umsteigen2}} \\ \beta_{\text{Umsteigen3}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Pünktlichkeit2}} \\ x_{\text{Pünktlichkeit3}} \\ x_{\text{Takt2}} \\ x_{\text{Takt3}} \\ x_{\text{Umsteigen2}} \\ x_{\text{Umsteigen3}} \end{bmatrix} + \begin{bmatrix} \gamma_{\text{ZeitÖV}} \\ \gamma_{\text{PreisÖV}} \\ \gamma_{\text{Komfort}} \\ \gamma_{\text{Inform.}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Zeit}} \\ x_{\text{Preis}} \\ x_{\text{Komfort}} \\ x_{\text{Inform.}} \end{bmatrix} \quad (7.2)$$

sowie

$$V(\text{Pkw}) = \begin{bmatrix} \gamma_{\text{Pkw}} \\ \gamma_{\text{ZeitIV}} \\ \gamma_{\text{KostenIV}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Pkw}} \\ x_{\text{Zeit}} \\ x_{\text{Kosten}} \end{bmatrix}. \quad (7.3)$$

Wie in (4.4) erfolgte bei den Nutzenfunktionen der Regionalbahn und des Regionalbusses eine Trennung zwischen den Nutzenparametern β , welche sich auf die Attribute des zerlegten Konstrukts beziehen, und den Nutzenparametern γ , welche sich auf die beiden übrigen Konstrukte, die alternativenspezifische Konstante sowie die Zeit- und Preisattribute beziehen. Für die beiden öffentlichen Verkehrsmittel wurden alternativenüber-

⁴⁰⁴ Die Schätzung dieser und der folgenden Modelle erfolgte mit *Stata 12*, vgl. *StataCorp.* (2011). Zur Darstellung der Ergebnisse wurde das zugehörige User-Programm *estout* verwendet, vgl. *Jann* (2005 und 2007).

greifende Parameter für die Zeit-, Preis- und Qualitätsattribute sowie für die Konstrukte geschätzt.⁴⁰⁵ Für die Pkw-Alternative wurden alternativenspezifische Parameter für die Zeit- und Kostenattribute geschätzt.

Die Ergebnisse der geschätzten multinomialen Logit-Modelle für die Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* sind Tab. 7.8, und für die Gruppen der *Pkw-Nutzer* und der „*wahlfreien*“ *Pkw-Nutzer* in Tab. 7.9 dargestellt. Die Tabellen enthalten jeweils die Schätzwerte der Nutzenparameter β bzw. γ und deren zugehörige *t*-Statistiken.

Tab. 7.8: *Multinomiales Logit-Modell der Verbindungsqualität für die nicht wahlfreien und die wahlfreien Bahnnutzer*

| | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | | |
|--|--|-----------|-------------------------|-----------|---------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) | |
| Bahnkonstante | 0,497 | (7,20) | 0,840 | (9,67) | |
| Pkw-Konstante | | | 1,429 | (4,08) | |
| Zeit ÖV | -0,055 | (-7,54) | -0,033 | (-7,85) | |
| Preis ÖV | -0,674 | (-4,64) | -0,182 | (-6,98) | |
| Zeit IV | | | -0,036 | (-6,48) | |
| Kosten IV | | | -0,137 | (-6,02) | |
| Komfort | 0,187 | (7,41) | 0,098 | (3,82) | |
| Informationen | 0,086 | (3,79) | 0,076 | (2,94) | |
| Pünktlichkeit | | | | | |
| | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,321 | 0,263 | | |
| | 2. <i>3 bis 10 Minuten Verspätung</i> | 0,072 | (1,04) | 0,121 | (1,65) |
| | 3. <i>10 bis 20 Minuten Verspätung</i> | -0,394 | (-5,70) | -0,384 | (-5,03) |
| Umsteigehäufigkeit | | | | | |
| | 1. <i>0 x Umsteigen</i> | 0,403 | 0,357 | | |
| | 2. <i>1 x Umsteigen mit Anschlussgarantie</i> | -0,117 | (-1,67) | -0,027 | (-0,36) |
| | 3. <i>1 x Umsteigen ohne Anschlussgarantie</i> | -0,286 | (-4,16) | -0,330 | (-4,28) |
| Takt | | | | | |
| | 1. <i>alle 30 Minuten</i> | 0,683 | 0,621 | | |
| | 2. <i>alle 60 Minuten</i> | 0,018 | (0,27) | 0,158 | (2,15) |
| | 3. <i>alle 120 Minuten</i> | -0,702 | (-9,64) | -0,779 | (-9,61) |
| Anzahl Wahlsituationen | 1213 | | 1010 | | |
| Anzahl Parameter | 11 | | 14 | | |
| Log-Likelihood-Wert des Modells | -654,25 | | -880,09 | | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | -820,54 | | -1052,81 | | |
| Log-Likelihood-Wert des zugeh. Nullmodells | -840,79 | | -1109,60 | | |

Quelle: *Eigene Berechnungen.*

⁴⁰⁵ Eine Überprüfung zeigte, dass alternativenspezifische Parameter für die beiden ÖV-Alternativen nicht zu signifikanten Verbesserungen dieser und der folgenden Modelle führten.

Tab. 7.9: *Multinomiales Logit-Modell der Verbindungsqualität für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer*

| | | Pkw-Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|--|--|------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,251 | (3,39) | 0,269 | (3,35) |
| Pkw-Konstante | | 2,057 | (9,57) | 1,329 | (5,21) |
| Zeit ÖV | | -0,050 | (-11,69) | -0,069 | (-12,77) |
| Preis ÖV | | -0,038 | (-0,77) | -0,171 | (-2,23) |
| Zeit IV | | -0,052 | (-10,46) | -0,075 | (-11,81) |
| Kosten IV | | -0,030 | (-1,10) | -0,138 | (-3,19) |
| Komfort | | 0,078 | (3,93) | 0,084 | (3,65) |
| Informationen | | -0,002 | (-0,09) | -0,003 | (-0,13) |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,243 | | 0,290 | |
| | 2. <i>3 bis 10 Minuten Verspätung</i> | 0,154 | (2,73) | 0,168 | (2,55) |
| | 3. <i>10 bis 20 Minuten Verspätung</i> | -0,397 | (-6,47) | -0,458 | (-6,53) |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | 0,394 | | 0,454 | |
| | 2. <i>1 x Umsteigen mit Anschlussgarantie</i> | -0,028 | (-0,49) | -0,036 | (-0,54) |
| | 3. <i>1 x Umsteigen ohne Anschlussgarantie</i> | -0,366 | (-5,97) | -0,418 | (-5,95) |
| Takt | 1. <i>alle 30 Minuten</i> | 0,380 | | 0,443 | |
| | 2. <i>alle 60 Minuten</i> | 0,063 | (1,12) | 0,100 | (1,51) |
| | 3. <i>alle 120 Minuten</i> | -0,443 | (-7,12) | -0,543 | (-7,50) |
| Anzahl Wahlsituationen | | 2519 | | 1338 | |
| Anzahl Parameter | | 14 | | 14 | |
| Log-Likelihood-Wert des Modells | | -1947,84 | | -1234,22 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | | -2119,16 | | -1432,40 | |
| Log-Likelihood-Wert des zugeh. Nullmodells | | -2767,40 | | -1469,94 | |

Quelle: Eigene Berechnungen.

Der Nutzenparameter der Referenzkategorie der effektkodierten Attribute entspricht der negativen Summe der Nutzenparameter der übrigen Attributsausprägungen.

Darüber hinaus enthalten Tab. 7.8 und Tab. 7.9 die Anzahl der Wahlentscheidungen, die Anzahl der Parameter und den Log-Likelihood-Wert des MNL-Modells sowie zum Vergleich die Log-Likelihood-Werte des zugehörigen Konstanten- und des Nullmodells. Im Konstantenmodell werden durch die alternativenspezifischen Konstanten die aggregierten Wahlanteile der Alternativen exakt wiedergegeben. Dahingegen wird

beim Nullmodell eine vollkommen zufällige Wahl unterstellt, sodass die Wahlanteile aller Alternativen gleich groß sind.⁴⁰⁶

Wie in Kap. 3.5.2 und Kap. 3.5.3 beschrieben erfolgte die Überprüfung der Gesamtsignifikanz des jeweiligen MNL-Modells durch einen Vergleich des MNL-Modells (unrestringiertes Modell) mit dem Konstantenmodell (restringiertes Modell)⁴⁰⁷ anhand des Likelihood-Ratio-Tests. Die dazu verwendete Teststatistik lautet $\lambda = -2[LL_{\text{Konstantenmodell}} - LL_{\text{MNL}}]$. Sofern das Testergebnis den kritischen Wert χ_{krit}^2 überschreitet, wird das Konstantenmodell zugunsten des MNL-Modells verworfen. Die Ergebnisse der Überprüfung der Gesamtsignifikanz sind in Tab. 7.10 dargestellt. Bei allen Nutzergruppen überschreitet der Wert der Teststatistik den jeweiligen kritischen Wert. Das Konstantenmodell wurde somit bei jeder der vier Gruppen zugunsten des MNL-Modells verworfen.

Tab. 7.10: Überprüfung der Gesamtsignifikanz des MNL-Modells der Verbindungsqualität differenziert nach Nutzergruppen

| Nutzergruppe | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ_{krit}^2 ($\alpha = 0,05$) |
|----------------------------|--|---------------------------|-------------------------------------|
| Nicht wahlfreie Bahnnutzer | $\lambda = -2[-820,54 - (-654,25)] = 332,58$ | 11-1=10 | 18,31 |
| Wahlfreie Bahnnutzer | $\lambda = -2[-1052,81 - (-880,09)] = 345,43$ | 14-2=12 | 21,03 |
| Pkw-Nutzer | $\lambda = -2[-2119,16 - (-1947,84)] = 342,65$ | 14-2=12 | 21,03 |
| „Wahlfreie“ Pkw-Nutzer | $\lambda = -2[-1432,40 - (-1234,22)] = 396,36$ | 14-2=12 | 21,03 |

Quelle: Eigene Berechnungen.

Allgemein folgt aus der Gesamtsignifikanz des MNL-Modells nicht notwendigerweise, dass auch alle einzelnen Nutzenparameter signifikant von 0 verschieden sind. Die Teststatistik für die Signifikanz der einzelnen Parameter stellt die t -Statistik dar. Die t -Statistik wird aus dem Quotienten des Nutzenparameters und dessen Standardfehler berechnet. Sofern die t -Statistik den kritischen Wert $t_{krit} = 1,96$ (Signifikanzniveau von $\alpha = 0,05$) betragsmäßig überschreitet, wird die Nullhypothese, dass der geschätzte Nutzenparameter einen Wert von 0 aufweist, verworfen. Ein nicht signifikant von 0 verschiedener Parameter bedeutet, dass keine Abweichung von der alternativenspezifischen Konstante besteht.

⁴⁰⁶ Vgl. Louviere et al. (2000), S. 274.

⁴⁰⁷ Beim Konstantenmodell besteht die Restriktion dahingehend, dass die Nutzenparameter auf 0 gesetzt werden, d. h. es werden keine Nutzenparameter geschätzt.

Für das Modell der Gruppe der *nicht wahlfreien Bahnnutzer* konnten bis auf wenige Ausnahmen durchgehend signifikante Parameter geschätzt werden. Die Ausnahmen bilden die Parameter, die sich auf die zweite von insgesamt drei Ausprägungen der Attribute *Pünktlichkeit*, *Umsteigen* und *Takt* beziehen. Da jedoch der Parameter, der sich jeweils auf die dritte Ausprägung bezieht, signifikant ist, wurde gefolgert, dass sich der Nutzen linear mit den Ausprägungen der Attribute *Pünktlichkeit*, *Umsteigen* und *Takt* verändert. Das bedeutet, dass die Veränderung des Nutzens von der ersten zur zweiten Ausprägungen genauso groß ist wie die Veränderung des Nutzens von der zweiten zur dritten Ausprägung; allerdings weisen die Attributsausprägungen an sich keinen linearen Verlauf auf.⁴⁰⁸

Alle Parameter weisen plausible Vorzeichen auf: Mit einer zunehmenden *Gesamtreisezeit* sowie mit einem zunehmenden *Fahrpreis* sinkt der Nutzen des Verkehrsmittels (negatives Vorzeichen), wohingegen mit höheren zusammenfassenden Konstruktwerten für den *Komfort* und die *Informationen* sowie mit besseren Attributsausprägungen für *Pünktlichkeit*, *Umsteigen* und *Takt* der Nutzen zunimmt (positives Vorzeichen). Der Parameter des Konstrukts des *Komforts* weist einen größeren Wert auf als der Parameter des Konstrukts *Informationen*. Die positive Bahnkonstante bedeutet, dass die Regionalbahn bei gleichen Ausprägungen der Attribute und Konstrukte einen größeren Nutzen als der Regionalbus stiftet.

Von den Parametern, die für das Modell der Gruppe der *wahlfreien Bahnnutzer* geschätzt wurden, sind alle bis auf die Parameter, die sich auf die zweite von insgesamt drei Ausprägungen der Attribute *Pünktlichkeit* und *Umsteigen* beziehen, signifikant. Auch hier lässt sich folgern, dass sich der Nutzen linear mit den Ausprägungen dieser Attribute verändert. Zudem weisen alle Parameter plausible Vorzeichen auf. Wie bei der vorherigen Gruppe weist der Parameter des Konstrukts des *Komforts* einen größeren Wert auf als der Parameter des Konstrukts der *Informationen* und es liegt eine positive Bahnkonstante vor.

Die Pkw-Konstante ist ebenfalls positiv. Bezüglich dieser Konstante kann zunächst kein direkter Vergleich bei „ansonsten gleichen Ausprägungen“

⁴⁰⁸ Beispielsweise wird beim *Takt* zwischen einem 30-, 60- und 120-Minutentakt unterschieden.

erfolgen, da die Nutzenfunktion wie in (7.3) dargestellt nicht die entsprechenden Nutzenparameter und effektkodierten Attribute enthält. Ein Interpretation der Konstanten ist jedoch folgendermaßen möglich: Die (effektkodierten) Attribute und Konstrukte der ÖV-Alternativen mit den zugehörigen Nutzenparametern können der Nutzenfunktion des Pkw hinzugefügt werden; da diese (kodierten) Attribute und Konstrukte durchgehend den Wert 0 aufweisen, verändert sich die Nutzenfunktion nicht. Ebenso können in die Nutzenfunktionen der öffentlichen Verkehrsmittel die Kosten- und Zeitattribute und Nutzenparameter des Pkw aufgenommen werden, ohne dass sich die Nutzenfunktionen verändern. Gleiche Ausprägungen der qualitativen Attribute und Konstrukte zur Interpretation der Konstanten liegen somit nur im hypothetischen Fall vor, wenn die Ausprägungen hierfür auch bei den öffentlichen Verkehrsmitteln 0 betragen. Bei der Referenzalternative sind Ausprägungen der effektkodierten Attribute von 0 mit dem durchschnittlichen, durch die qualitativen Attribute gestifteten Nutzen gleichzusetzen. In Bezug auf die Interpretation der positiven Pkw-Konstante bedeutet dies, dass der Pkw unter den genannten Bedingungen einen größeren Nutzen als der Regionalbus mit durchschnittlichen qualitativen Attributen stiftet.

In dem Modell der Gruppe der *Pkw-Nutzer* sind alle bis auf fünf Parameter signifikant. Zum einen handelt es sich dabei um Parameter, die sich auf die zweite von insgesamt drei Ausprägungen der Attribute *Umsteigen* und *Takt* beziehen. Daher wurde in Analogie zu den anderen Gruppen auch hier gefolgert, dass sich der Nutzen linear mit den Ausprägungen dieser Attribute verändert. Zum anderen sind der ÖV-spezifische Preisparameter, der IV-spezifische Kostenparameter und der Parameter des Konstrukts der *Informationen* nicht signifikant. Diese nicht signifikanten Parameter deuten darauf hin, dass die beiden Attribute und das Konstrukt von den Befragten nicht bzw. nicht gleichermaßen in die Wahlentscheidung einbezogen wurden.

Bei der Gruppe der *Pkw-Nutzer* fällt zudem auf, dass es einen deutlichen Unterschied zwischen dem Log-Likelihood-Wert des Konstantenmodells

und dem des Nullmodells gibt.⁴⁰⁹ Dies kann als Indikator dafür angesehen werden, dass sich die Wahlanteile der Alternativen stark unterscheiden. Eine Auswertung der aggregierten Wahlanteile ergibt einen Anteil von 68,4 % für den Pkw, von 17,7 % für die Regionalbahn und von 14,0 % für den Regionalbus. Aufgrund dieses relativ großen Anteils der Pkw-Alternative stellte sich die Frage, ob die *Pkw-Nutzer* insgesamt die beiden öffentlichen Verkehrsmittel selten gewählt haben oder ob es eine Gruppe von *Pkw-Nutzern* gibt, die in keiner der Wahlsituationen zwischen den Alternativen gewechselt hat.

Eine Analyse der individuellen Wahlentscheidungen zeigte, dass 49,3 % der *Pkw-Nutzer* bei den Wahlexperimenten dieses Subexperiments stets die gleiche Alternative gewählt haben.⁴¹⁰ Dieser Gruppe sind 46,9 % der Wahlsituationen zuzuordnen; in 96,0 % der Wahlsituationen, in denen nicht gewechselt wurde, wurde der Pkw gewählt.

Diese Teilgruppe der *Pkw-Nutzer* hat bei den gegebenen Wahlexperimenten nicht auf die Veränderungen der Ausprägungen der Attribute und Konstrukte reagiert. Möglicherweise waren für diese Teilgruppe die Veränderungen nicht groß genug. Bei den Experimenten wurde jedoch versucht, die Ausprägungen möglichst realistisch auszuwählen, damit die Alternativen als glaubhaft wahrgenommen werden. Durch eine deutlichere Verbesserung der beiden öffentlichen Verkehrsmittel und/oder eine deutlichere Verschlechterung der Pkw-Alternative hätte möglicherweise der Anteil der nicht wechselnden *Pkw-Nutzer* reduziert werden können. Durch eine derartige Veränderung wären die Experimente jedoch weniger realistisch und weniger glaubhaft geworden, wodurch gegebenenfalls insgesamt die Wahlexperimente nicht ernst genommen worden wären und die Bereitschaft zur Fortsetzung der Experimente gesunken wäre.

Zum Vergleich kann die Studie „Mobilität in Deutschland 2008“ herangezogen werden. Im Rahmen dieser Studie wurde eine Segmentierung von „mobilen Personen“ anhand der Verkehrsmittelnutzung und Erreich-

⁴⁰⁹ Der Likelihood-Ratio-Test zum Vergleich des Konstantenmodells mit dem Nullmodell führt zu einem Testergebnis von $-2[-2767,40 - (-2119,16)] = 1296,48$ bei einem kritischen Wert $\chi^2_{krit} = 5,99$ (χ^2 -Verteilung mit $2-0=2$ Freiheitsgraden).

⁴¹⁰ Bei den Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* ist dieser Anteil deutlich geringer.

barkeit der Ziele mit dem ÖPNV vorgenommen. Von den Personen, die (potenziell) über einen Pkw verfügen und die den ÖPNV *selten* oder *nie* nutzen, verfügen 35 % über eine nach eigener Einschätzung *gute* oder *sehr gute Erreichbarkeit der Ziele* mit dem ÖPNV.⁴¹¹ Letztere Gruppe umfasst 23 % der Personen mit Pkw.

Die Gründe für das (Nicht-)Nutzungsverhalten können unterschiedlich sein. Grundsätzlich wird die Verkehrsmittelwahl nicht nur durch das Verkehrsangebot, sondern auch durch personenbezogene Merkmale beeinflusst.⁴¹² Diese personenbezogenen Merkmale schränken die Pkw-Nutzer in ihrer Wahlfreiheit ein. Vor dem Hintergrund der Daten von „Mobilität in Deutschland“ erscheint der Anteil der *Pkw-Nutzer*, die im Wahlexperiment nicht zwischen den Alternativen gewechselt haben, plausibel.

Im beschriebenen Wahlmodell spiegelt sich der große Anteil der Befragten, die unabhängig von den Ausprägungen der Attribute und Konstrukte durchgehend den Pkw gewählt haben, in nicht signifikanten Nutzenparametern und einem hohen Erklärungsgehalt durch die Konstanten wieder. Für die Gruppe der *Pkw-Nutzer*, die hingegen auf Veränderungen der Ausprägungen von Attributen und Konstrukten reagiert haben, da sie zwischen den Alternativen gewechselt haben, wurde ein eigenes multinomiales Logit-Modell geschätzt. Die Befragten dieser Teilgruppe werden im Folgenden als „*wahlfreie*“ *Pkw-Nutzer* bezeichnet.

Von den Parametern, die für das Modell der Gruppe der „*wahlfreien*“ *Pkw-Nutzer* geschätzt wurden, sind alle bis auf drei signifikant. Bei den nicht signifikanten Parametern handelt es sich zum einen um Parameter, die sich auf die zweite von insgesamt drei Ausprägungen der Attribute *Umsteigen* und *Takt* beziehen, sodass ein linearer Zusammenhang gefolgert wurde. Zum anderen handelt es sich um den Parameter, der sich auf das Konstrukt der *Informationen* bezieht. Für die beiden Parameter, die sich auf die Preis- bzw. Kostenattribute beziehen, konnten im Gegensatz zu denen im Modell für alle *Pkw-Nutzer* signifikante Werte geschätzt werden.

⁴¹¹ Vgl. Lenz et al. (2010), S. 111–112.

⁴¹² Vgl. Boltze et al. (2002), S. 29.

Die Parameter weisen plausible Vorzeichen auf. Sowohl die Bahn- als auch die Pkw-Konstante sind positiv, d. h. die Regionalbahn und der Pkw stiften bei ansonsten gleichen Ausprägungen einen größeren Nutzen als der Regionalbus.

7.5.2 Subexperiment des Komforts

Differenziert nach Nutzergruppen wurden für das Subexperiment des *Komforts* multinomiale Logit-Modelle geschätzt. Die Nutzenfunktion der Regionalbahn, die diesen Modellen zugrunde liegt, lautet

$$V(\text{Bahn}) = \begin{bmatrix} \beta_{\text{Toilette2}} \\ \beta_{\text{Toilette3}} \\ \beta_{\text{SauberkeitInnen2}} \\ \beta_{\text{Sitzplatzverfügbarkeit2}} \\ \beta_{\text{Sitzplatzverfügbarkeit3}} \\ \beta_{\text{Komfort Sitzplatz2}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Toilette2}} \\ x_{\text{Toilette3}} \\ x_{\text{SauberkeitInnen2}} \\ x_{\text{Sitzplatzverfügbarkeit2}} \\ x_{\text{Sitzplatzverfügbarkeit3}} \\ x_{\text{Komfort Sitzplatz2}} \end{bmatrix} + \quad (7.4)$$

$$\begin{bmatrix} \gamma_{\text{Bahn}} \\ \gamma_{\text{ZeitÖV}} \\ \gamma_{\text{PreisÖV}} \\ \gamma_{\text{Verbindungsqualität}} \\ \gamma_{\text{Informationen}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Bahn}} \\ x_{\text{Zeit}} \\ x_{\text{Preis}} \\ x_{\text{Verbindungsqualität}} \\ x_{\text{Informationen}} \end{bmatrix}.$$

Die Nutzenfunktion des Regionalbusses unterscheidet sich von der Nutzenfunktion der Regionalbahn in (7.4) dahingehend, dass keine alternativenspezifische Konstante vorhanden ist. Die Nutzenfunktion des Pkw ist unabhängig vom Subexperiment und entspricht (7.3).

In Analogie zum Subexperiment der *Verbindungsqualität* wurden auch bei diesem Subexperiment für die beiden öffentlichen Verkehrsmittel gemeinsame Parameter für die Zeit-, Preis- und Qualitätsattribute sowie für die Konstrukte geschätzt und für die Pkw-Alternative alternativenspezifische Parameter. Die Ergebnisse der geschätzten multinomialen Logit-Modelle für die Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* sowie der *Pkw-Nutzer* und der „*wahlfreien*“ *Pkw-Nutzer* sind in Tab. 7.11 bzw. Tab. 7.12 dargestellt.

Tab. 7.11: *Multinomiales Logit-Modell des Komforts für die nicht wahlfreien und die wahlfreien Bahnnutzer*

| | | | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | |
|---|----|--|-------------------------------|-----------|-------------------------|-----------|
| | | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | | 0,643 | (9,58) | 0,709 | (8,02) |
| Pkw-Konstante | | | | | 1,474 | (3,99) |
| Zeit ÖV | | | -0,073 | (-10,58) | -0,037 | (-8,49) |
| Preis ÖV | | | -0,829 | (-4,78) | -0,177 | (-6,15) |
| Zeit IV | | | | | -0,034 | (-6,18) |
| Kosten IV | | | | | -0,147 | (-6,42) |
| Verbindungsqualität Informationen | | | 0,177 | (7,53) | 0,134 | (5,01) |
| Toilette im Zug/Bus | | | | | 0,311 | |
| | 1. | <i>sauber</i> | 0,139 | | | |
| | 2. | verschmutzt | 0,029 | (0,44) | -0,171 | (-2,26) |
| | 3. | nicht vorhanden | -0,168 | (-2,59) | -0,140 | (-1,80) |
| Sauberkeit Zug/Bus innen | | | | | 0,283 | |
| | 1. | <i>sauber</i> | 0,128 | | | |
| | 2. | Boden klebrig, Papier liegt auf den Sitzen | -0,128 | (-2,78) | -0,283 | (-5,19) |
| Sitzplatzverfügbarkeit im Zug/Bus | | | | | 0,617 | |
| | 1. | <i>ganze Fahrt Sitzplatz</i> | 0,635 | | | |
| | 2. | halbe Fahrt Sitzplatz | 0,086 | (1,34) | -0,043 | (-0,57) |
| | 3. | kein freier Sitzplatz | -0,721 | (-10,32) | -0,574 | (-7,22) |
| Komfort der Sitzplätze | | | | | 0,046 | |
| | 1. | <i>bequem</i> | 0,024 | | | |
| | 2. | unbequem | -0,024 | (-0,51) | -0,046 | (-0,86) |
| Anzahl Wahlsituationen | | | 1351 | | 945 | |
| Anzahl Parameter | | | 11 | | 14 | |
| Log-Likelihood-Wert des Modells | | | -707,23 | | -851,69 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | | | -894,98 | | -1002,54 | |
| Log-Likelihood-Wert des zugeh. Nullmodells | | | -936,44 | | -1038,19 | |

Quelle: *Eigene Berechnungen.*

Die Ergebnisse der Überprüfung der Gesamtsignifikanz der MNL-Modelle anhand von Likelihood-Ratio-Tests sind in Tab. 7.13 dargestellt. Für alle Gruppen wurde das Konstantenmodell zugunsten des MNL-Modells verworfen, da das jeweilige Testergebnis den kritischen Wert überschreitet.

Bezüglich der Signifikanz der einzelnen Nutzenparameter zeigen die jeweiligen *t*-Statistiken, dass für das Modell der Gruppe der *nicht wahlfreien Bahnnutzer* bis auf wenige Ausnahmen durchgehend von 0 signifikant verschiedene Parameter geschätzt wurden. Die Ausnahmen bilden

Tab. 7.12: *Multinomiales Logit-Modell des Komforts für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer*

| | Pkw-Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|---|------------|-----------|---------------------------|-----------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | 0,349 | (4,68) | 0,415 | (5,11) |
| Pkw-Konstante | 3,144 | (13,60) | 2,660 | (9,84) |
| Zeit ÖV | -0,059 | (-13,41) | -0,081 | (-14,76) |
| Preis ÖV | -0,231 | (-4,20) | -0,164 | (-2,33) |
| Zeit IV | -0,060 | (-11,76) | -0,085 | (-13,02) |
| Kosten IV | -0,142 | (-4,70) | -0,134 | (-3,46) |
| Verbindungsqualität Informationen | 0,168 | (8,09) | 0,196 | (8,18) |
| Toilette im Zug/Bus | | | | |
| 1. sauber | 0,169 | | 0,212 | |
| 2. verschmutzt | -0,119 | (-2,04) | -0,180 | (-2,68) |
| 3. nicht vorhanden | -0,049 | (-0,86) | -0,033 | (-0,49) |
| Sauberkeit Zug/Bus innen | | | | |
| 1. sauber | 0,251 | | 0,336 | |
| 2. Boden klebrig, Papier liegt auf den Sitzen | -0,251 | (-6,13) | -0,336 | (-7,17) |
| Sitzplatzverfügbar- keit im Zug/Bus | | | | |
| 1. ganze Fahrt Sitzplatz | 0,302 | | 0,381 | |
| 2. halbe Fahrt Sitzplatz | 0,082 | (1,44) | 0,069 | (1,06) |
| 3. kein freier Sitzplatz | -0,384 | (-6,19) | -0,450 | (-6,39) |
| Komfort der Sitzplätze | | | | |
| 1. bequem | 0,166 | | 0,234 | |
| 2. unbequem | -0,166 | (-4,07) | -0,234 | (-4,94) |
| Anzahl Wahlsituationen | 2455 | | 1422 | |
| Anzahl Parameter | 14 | | 14 | |
| Log-Likelihood-Wert des Modells | -1899,19 | | -1263,01 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | -2105,58 | | -1505,61 | |
| Log-Likelihood-Wert des zugeh. Nullmodells | -2697,09 | | -1562,23 | |

Quelle: Eigene Berechnungen.

Tab. 7.13: *Überprüfung der Gesamtsignifikanz des MNL-Modells des Komforts differenziert nach Nutzergruppen*

| Nutzergruppe | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheits- grade | χ^2_{krit} ($\alpha = 0,05$) |
|----------------------------|--|-----------------------------------|--|
| Nicht wahlfreie Bahnnutzer | $\lambda = -2[-894,98 - (-707,23)] = 375,48$ | 11-1=10 | 18,31 |
| Wahlfreie Bahnnutzer | $\lambda = -2[-1002,54 - (-851,69)] = 301,70$ | 14-2=12 | 21,03 |
| Pkw-Nutzer | $\lambda = -2[-2105,58 - (-1899,19)] = 412,78$ | 14-2=12 | 21,03 |
| „Wahlfreie“ Pkw-Nutzer | $\lambda = -2[-1505,61 - (-1263,01)] = 485,19$ | 14-2=12 | 21,03 |

Quelle: Eigene Berechnungen.

zum einen Parameter, die sich auf die zweite von insgesamt drei Ausprägungen der Attribute *Toilette im Zug/Bus* und *Sitzplatzverfügbarkeit im*

Zug/Bus beziehen, sodass erneut gefolgert wurde, dass sich der Nutzen linear mit den Ausprägungen dieser Attribute verändert. Zum anderen weist der Parameter des Attributs *Komfort der Sitzplätze* keinen signifikant von 0 verschiedenen Wert auf. Da vermutet wurde, dass dieses Attribut nur dann von Bedeutung ist, wenn ein freier Sitzplatz vorhanden ist, wurde ein Modell mit einem zusätzlichen Parameter für die Interaktion zwischen der *Sitzplatzverfügbarkeit* und dem *Komfort der Sitzplätze* geschätzt. Da sich dieser Parameter jedoch als nicht signifikant erwies, wurde das entsprechende Modell verworfen.

Die signifikanten Parameter der Attribute und Konstrukte weisen plausible Vorzeichen auf. Der Parameter des Konstrukts der *Verbindungsqualität* ist wertmäßig größer als der Parameter des Konstrukts der *Informationen*. Aufgrund der positiven Bahnkonstante wurde gefolgert, dass die Regionalbahn bei ansonsten gleichen Ausprägungen einen größeren Nutzen als der Regionalbus stiftet.

Für das Modell der Gruppe der *wahlfreien Bahnnutzer* wurden bis auf drei Ausnahmen durchgehend signifikante Parameter geschätzt. Bei diesen Ausnahmen handelt es sich zum Ersten um den Parameter, der sich auf die dritte Ausprägung des Attributs *Toilette im Zug/Bus* bezieht. Da für die zweite Ausprägung hingegen ein signifikanter Parameter geschätzt wurde, dessen Wert kleiner als 0 ist, wurde gefolgert, dass sich der Nutzen linear mit den Ausprägungen des Attributs verändert, jedoch in einer anderen Reihenfolge als während der Effektkodierung angenommen, d. h. eine *nicht vorhandene Toilette* (dritte Ausprägung) stiftet einen größeren Nutzen als eine *verschmutzte Toilette* (zweite Ausprägung). Die zweite Ausnahme stellt der Parameter dar, der sich auf die zweite von insgesamt drei Ausprägungen des Attributs *Sitzplatzverfügbarkeit* bezieht, sodass deswegen erneut gefolgert wurde, dass sich der Nutzen linear mit den Ausprägungen dieses Attributs verändert. Die dritte Ausnahme stellt der Parameter des Attributs *Komfort der Sitzplätze* dar. Ein Modell mit einem zusätzlichen Parameter für die Interaktion zwischen *Sitzplatzverfügbarkeit* und *Komfort der Sitzplätze* wurde verworfen, da sich ebendieser Parameter als nicht signifikant erwies.

Die signifikanten Parameter der Attribute und Konstrukte weisen plausible Vorzeichen auf. Der Parameter des Konstrukts der *Verbindungs-*

qualität ist wertmäßig größer als der Parameter des Konstrukts der *Informationen* und sowohl die Bahnkonstante als auch die Pkw-Konstante weisen positive Werte auf.

Für das Modell der Gruppe der *Pkw-Nutzer* wurden bis auf den Parameter, der sich auf die dritte Ausprägung des Attributs *Toilette im Zug/Bus* bezieht und den Parameter, der sich auf die zweite Ausprägung des Attributs *Sitzplatzverfügbarkeit im Zug/Bus* bezieht, durchgehend signifikante Parameter geschätzt. In Kombination mit den übrigen Schätzergebnissen wurde gefolgert, dass sich der Nutzen linear mit den Ausprägungen des Attributs *Toilette im Zug/Bus* bzw. *Sitzplatzverfügbarkeit im Zug/Bus* ändert. Genau wie bei der Gruppe der *wahlfreien Bahnnutzer* weicht auch hier die Reihenfolge der Ausprägungen des Attributs *Toilette im Zug/Bus* von der während der Effektkodierung angenommenen Reihenfolge ab.

Erneut weisen die signifikanten Parameter der Attribute und Konstrukte plausible Vorzeichen auf. Auch im Modell dieser Gruppe ist der Parameter des Konstrukts der *Verbindungsqualität* wertmäßig größer als der Parameter des Konstrukts der *Informationen* und die alternativenspezifischen Konstanten weisen positive Werte auf, sodass deren Interpretation in Analogie zu den anderen Modellen erfolgt.

Im Gegensatz zu den Modellergebnissen für das Subexperiment der *Verbindungsqualität* besteht bei diesem Modell nicht das Problem der nicht signifikanten Parameter. Jedoch deutet auch hier der Unterschied zwischen dem Log-Likelihood-Wert des Konstantenmodells und des Nullmodells⁴¹³ darauf hin, dass sich die Wahlanteile der Alternativen deutlich unterscheiden. Eine Analyse der aggregierten Wahlanteile bestätigt dies: Der Pkw wurde in 67,1 %, die Regionalbahn in 19,0 % und der Regionalbus in 13,9 % der Wahlsituationen gewählt. Eine Analyse der individuellen Wahlentscheidungen zeigt, dass in 44,8 % der Wahlsituationen bzw. von 42,1 % der *Pkw-Nutzer* stets die gleiche Alternative gewählt wurde.⁴¹⁴ In 96,8 % der Wahlsituationen, in denen nicht gewechselt wurde, wurde ein Pkw gewählt.

⁴¹³ Der Likelihood-Ratio-Test zum Vergleich des Konstantenmodells mit dem Nullmodell führt zu einem Testergebnis von $\lambda = -2[-2697,09 - (-2105,58)] = 1183,03$ bei einem kritischen Wert $\chi^2_{krit} = 5,99$ (χ^2 -Verteilung mit $2-0=2$ Freiheitsgraden).

⁴¹⁴ Bei den Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* sind diese Anteile deutlich geringer.

Die Schätzergebnisse für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* für das Subexperiment des *Komforts* zeigen hinsichtlich der Signifikanz der Nutzenparameter, der Vorzeichen der Parameter, der alternativenspezifischen Konstanten sowie den daraus resultierenden Interpretationen keine Veränderungen im Vergleich zum Modell der Gruppe aller *Pkw-Nutzer*.

7.5.3 Subexperiment der Informationen

Auch für das Subexperiment der *Informationen* wurden differenziert nach Nutzergruppen multinomiale Logit-Modelle geschätzt. Die diesen Modellen zugrunde liegende Nutzenfunktion der Regionalbahn ist in (7.5) dargestellt. Die Nutzenfunktion des Regionalbusses unterscheidet sich von der Nutzenfunktion der Regionalbahn in (7.5) durch die fehlende alternativenspezifische Konstante; die Nutzenfunktion des Pkw entspricht (7.3).

$$V(\text{Bahn}) =$$

$$\begin{bmatrix} \beta_{\text{Fahrplaninfo2}} \\ \beta_{\text{Durchsage Störungen im Zug/Bus}} \\ \beta_{\text{Anzeige Störungen im Zug/Bus}} \\ \beta_{\text{Interaktion Störungen im Zug/Bus}} \\ \beta_{\text{Durchsage Störungen Bahnhof/Bush.}} \\ \beta_{\text{Anzeige Störungen Bahnhof/Bush.}} \\ \beta_{\text{Interaktion Störungen Bahnhof/Bush.}} \\ \beta_{\text{Durchsage Anschluss im Zug/Bus}} \\ \beta_{\text{Anzeige Anschluss im Zug/Bus}} \\ \beta_{\text{Interaktion Anschluss im Zug/Bus}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Fahrplaninfo2}} \\ x_{\text{Durchsage Störungen im Zug}} \\ x_{\text{Anzeige Störungen im Zug}} \\ x_{\text{Interaktion Störungen im Zug}} \\ x_{\text{Durchsage Störungen Bahnhof}} \\ x_{\text{Anzeige Störungen Bahnhof}} \\ x_{\text{Interaktion Störungen Bahnhof}} \\ x_{\text{Durchsage Anschluss im Zug}} \\ x_{\text{Anzeige Anschluss im Zug}} \\ x_{\text{Interaktion Anschluss im Zug}} \end{bmatrix} + \quad (7.5)$$

$$\begin{bmatrix} \gamma_{\text{Bahn}} \\ \gamma_{\text{ZeitÖV}} \\ \gamma_{\text{PreisÖV}} \\ \gamma_{\text{Verbindungsqualität}} \\ \gamma_{\text{Komfort}} \end{bmatrix}' \cdot \begin{bmatrix} x_{\text{Bahn}} \\ x_{\text{Zeit}} \\ x_{\text{Preis}} \\ x_{\text{Verbindungsqualität}} \\ x_{\text{Komfort}} \end{bmatrix}$$

Wie bei den anderen Subexperimenten wurden für die beiden öffentlichen Verkehrsmittel alternativensübergreifende Parameter und für den Pkw alternativenspezifische Parameter geschätzt. Die Modellergebnisse sind für die vier Nutzergruppen in Tab. 7.14 und Tab. 7.15 dargestellt.

Tab. 7.14: *Multinomiales Logit-Modell der Informationen für die nicht wahlfreien und die wahlfreien Bahnnutzer*

| | Nicht wahlfreie Bahnnutzer | | Wahlfreie Bahnnutzer | | |
|---|--|--------------------------|-----------------------------|----------------------------|-------------------------------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) | |
| Bahnkonstante | 0,664 | (9,11) | 1,047 | (10,57) | |
| Pkw-Konstante | | | 1,518 | (3,38) | |
| Zeit ÖV | -0,074 | (-10,10) | -0,057 | (-9,59) | |
| Preis ÖV | -0,518 | (-3,26) | -0,244 | (-7,05) | |
| Zeit IV | | | -0,061 | (-7,46) | |
| Kosten IV | | | -0,140 | (-4,97) | |
| Verbindungsqualität | 0,242 | (9,27) | 0,185 | (6,05) | |
| Komfort | 0,156 | (6,31) | 0,116 | (3,82) | |
| Fahrplaninformationen | 1. <i>vorhanden</i> | 0,119 | 0,122 | | |
| am Bahnhof/ an Bushaltestelle | 2. <i>nicht vorhanden</i> | -0,119 | -0,122 | (-1,98) | |
| Informationen im Zug/Bus zu Grund und Dauer von Störungen | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | 0,023 0,071 0,001 | (0,45) (1,44) (0,02) | 0,097 -0,022 -0,018 | (1,56) (-0,35) (-0,29) |
| Informationen am Bahnhof/Bushaltestelle zu Grund und Dauer von Störungen | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | 0,050 0,137 -0,055 | (1,01) (2,74) (-1,08) | -0,021 0,101 0,035 | (-0,34) (1,61) (0,57) |
| Informationen im Zug/Bus zu aktuellen Anschlusszügen/-bussen | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | 0,020 0,034 -0,108 | (0,41) (0,68) (-2,23) | -0,097 -0,030 -0,038 | (-1,58) (-0,50) (-0,62) |
| Anzahl Wahlsituationen | | 1121 | | 760 | |
| Anzahl Parameter | | 15 | | 18 | |
| Log-Likelihood-Wert des Modells | | -610,74 | | -608,43 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | | -745,40 | | -742,47 | |
| Log-Likelihood-Wert des zugeh. Nullmodells | | -777,02 | | -834,95 | |

Quelle: *Eigene Berechnungen.*

Die Ergebnisse des Likelihood-Ratio-Tests zur Überprüfung der Gesamtsignifikanz der MNL-Modelle sind in Tab. 7.16 dargestellt. Die Testergebnisse überschreiten durchgehend die kritischen Werte, sodass für alle Gruppen die Gesamtsignifikanz des MNL-Modells gegeben ist.

Zwar sind die MNL-Modelle insgesamt statistisch signifikant, doch zeigen die *t*-Statistiken Unterschiede in Bezug auf die Signifikanz der einzelnen Nutzenparameter. Für die Parameter, die sich auf die alternativenspezifischen Konstanten, Gesamtreisezeiten, Preise/Kosten, Konstrukte und das effektkodierte Attribut *Fahrplaninformationen am Bahnsteig* beziehen, konnten in allen Modellen durchgehend signifikante

Tab. 7.15: *Multinomiales Logit-Modell der Informationen für die Pkw-Nutzer und die „wahlfreien“ Pkw-Nutzer*

| | Pkw-Nutzer | | „Wahlfreie“ Pkw-Nutzer | |
|---|---------------------------|----------------|---------------------------|-----------|
| | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | 0,164 | (2,42) | 0,210 | (2,77) |
| Pkw-Konstante | 2,971 | (14,34) | 2,377 | (9,21) |
| Zeit ÖV | -0,044 | (-11,08) | -0,084 | (-15,46) |
| Preis ÖV | -0,154 | (-2,90) | -0,181 | (-2,01) |
| Zeit IV | -0,065 | (-13,84) | -0,093 | (-14,44) |
| Kosten IV | -0,058 | (-2,13) | -0,147 | (-2,96) |
| Verbindungsqualität | 0,137 | (7,28) | 0,193 | (8,45) |
| Komfort | 0,063 | (3,44) | 0,106 | (4,73) |
| Fahrplaninformationen | 1. <i>vorhanden</i> | 0,157 | 0,220 | |
| am Bahnhof/ an Bushaltestelle | 2. <i>nicht vorhanden</i> | -0,157 (-4,20) | -0,220 (-4,81) | |
| Informationen im Zug/Bus | Durchsageneffekt | 0,088 (2,35) | 0,124 (2,70) | |
| zu Grund und Dauer von | Anzeigeneffekt | 0,090 (2,40) | 0,109 (2,42) | |
| Störungen | Interaktionseffekt | -0,038 (-1,01) | -0,062 (-1,37) | |
| Informationen am | Durchsageneffekt | 0,046 (1,23) | 0,100 (2,17) | |
| Bahnhof/Bushaltestelle | Anzeigeneffekt | 0,029 (0,78) | 0,025 (0,56) | |
| zu Grund und Dauer von | Interaktionseffekt | -0,039 (-1,03) | -0,047 (-1,03) | |
| Störungen | | | | |
| Informationen im Zug/Bus | Durchsageneffekt | -0,060 (-1,59) | -0,050 (-1,10) | |
| zu aktuellen | Anzeigeneffekt | -0,009 (-0,25) | 0,017 (0,38) | |
| Anschlusszügen/ -bussen | Interaktionseffekt | -0,008 (-0,22) | -0,012 (-0,25) | |
| Anzahl Wahlsituationen | 2533 | | 1374 | |
| Anzahl Parameter | 18 | | 18 | |
| Log-Likelihood-Wert des Modells | -2153,12 | | -1278,11 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | -2307,41 | | -1496,52 | |
| Log-Likelihood-Wert des zugeh. Nullmodells | -2782,78 | | -1509,49 | |

Quelle: *Eigene Berechnungen.*

Tab. 7.16: *Überprüfung der Gesamtsignifikanz des MNL-Modells der Informationen differenziert nach Nutzergruppen*

| Nutzergruppe | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheits- grade | χ^2_{krit} ($\alpha = 0,05$) |
|----------------------------|--|-----------------------------------|--|
| Nicht wahlfreie Bahnnutzer | $\lambda = -2[-745,40 - (-610,74)] = 269,31$ | 15-1=14 | 23,68 |
| Wahlfreie Bahnnutzer | $\lambda = -2[-742,47 - (-608,43)] = 268,08$ | 18-2=16 | 26,30 |
| Pkw-Nutzer | $\lambda = -2[-2307,41 - (-2153,12)] = 308,58$ | 18-2=16 | 26,30 |
| „Wahlfreie“ Pkw-Nutzer | $\lambda = 2[-1496,52 - (-1278,11)] = 436,80$ | 18-2=16 | 26,30 |

Quelle: *Eigene Berechnungen.*

Werte geschätzt werden. Zudem weisen diese Parameter plausible Vorzeichen auf.

Der Parameter des Konstrukts der *Verbindungsqualität* ist wertmäßig größer als der Parameter des Konstrukts des *Komforts*. Wie in den anderen Modellen wurden für die alternativenspezifischen Konstanten positive Werte geschätzt, sodass diesbezüglich mit den Modellen der anderen Subexperimente vergleichbare Ergebnisse vorliegen.

Für die anderen effektkodierten Attribute konnten hingegen nur vereinzelt signifikante Parameter geschätzt werden. Dabei handelt es sich um den *Anzeigeneffekt* der *Informationen am Bahnhof/an der Bushaltestelle zu Grund und Dauer von Störungen* und den *Interaktionseffekt* der *Informationen im Zug/im Bus zu aktuellen Anschlusszügen/-bussen* im Modell der *nicht wahlfreien Bahnnutzer*, den *Durchsagen-* und den *Anzeigeneffekt* der *Informationen im Zug/Bus zu Grund und Dauer von Störungen* in den Modellen der *Pkw-Nutzer* und der „*wahlfreien*“ *Pkw-Nutzer* sowie um den *Durchsageneneffekt* der *Informationen am Bahnhof/an der Bushaltestelle zu Grund und Dauer von Störungen* im Modell der „*wahlfreien*“ *Pkw-Nutzer*. Auch diese signifikanten Parameter weisen plausible Vorzeichen auf, d. h. übermittelte Informationen führen zu einem Nutzenzuwachs, wobei das negative Vorzeichen des Interaktionsparameters auf eine teilweise Redundanz der beiden Arten der Informationsvermittlung hinweist.

Auch bei diesem Subexperiment deutet im Modell der *Pkw-Nutzer* der Unterschied zwischen dem Log-Likelihood-Wert des Konstantenmodells und des Nullmodells auf deutlich ungleiche Wahlanteile hin.⁴¹⁵ Eine Analyse der aggregierten Wahlanteile zeigt, dass der Pkw in 63,3 %, die Regionalbahn in 20,0 % und der Regionalbus in 16,7 % der Wahlsituationen gewählt wurde. Zudem wurde von 48,8 % der *Pkw-Nutzer* bzw. in 45,8 % der Wahlsituationen stets die gleiche Alternative gewählt.⁴¹⁶ In 90,7 % dieser Situationen wurde ein Pkw gewählt.

⁴¹⁵ Der Likelihood-Ratio-Test zum Vergleich des Konstantenmodells mit dem Nullmodell führt zu einem Testergebnis von $\lambda = -2[-2782,78 - (-2307,41)] = 950,75$ bei einem kritischen Wert $\chi^2_{krit} = 5,99$ (χ^2 -Verteilung mit $2-0=2$ Freiheitsgraden).

⁴¹⁶ Bei den Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* sind diese Anteile deutlich geringer.

In Bezug auf mögliche Ursachen der nicht signifikanten Parameter sei zunächst auf die im Vergleich zu den anderen Subexperimenten größere Anzahl an Attributen und deren Ausprägungen mit zudem ausführlicheren Bezeichnungen hingewiesen. Darüber hinaus zeigten die Ergebnisse der anderen Subexperimente, dass das Konstrukt der *Informationen* insgesamt einen geringeren Nutzen als die Konstrukte der *Verbindungsqualität* und des *Komforts* stiftet.⁴¹⁷ Bei einer korrekten Abbildung der Attribute durch die Konstrukte würde daraus folgen, dass auch der Nutzen der einzelnen Attribute der *Informationen* geringer sein müsste. Möglicherweise ist dieser geringere Nutzen auch mit einer (teilweisen) Nicht-Beachtung der Attribute (*attribute non-attendance*)⁴¹⁸ oder einem nicht homogenen Wahlverhalten verbunden.

Aus Gründen der Vergleichbarkeit mit den anderen Modellen, in denen alle Informationsattribute indirekt durch den zusammenfassenden Konstruktwert enthalten sind, wurden die nicht signifikanten Parameter nicht aus dem Modell ausgeschlossen.

7.6 Überprüfung der Gleichheit der Nutzenfunktionen der Nutzergruppen I

Ausgehend von den multinomialen Logit-Modellen, die für die unterschiedlichen Nutzergruppen jeweils für die drei Subexperimente geschätzt wurden, wird im Folgenden überprüft, ob innerhalb eines jeden Subexperiments Gleichheit zwischen den Nutzenfunktionen der Nutzergruppen besteht. Sofern keine Unterschiede bestehen, können die Daten der Nutzergruppen eines Subexperiments gemeinsam ausgewertet werden.

Zur Überprüfung der Gleichheit wird ein zweistufiger Likelihood-Ratio-Test (vgl. 3.5.2) verwendet. Dabei werden zum einen die beiden Gruppen der *Bahnnutzer* miteinander verglichen, zum anderen die *wahlfreien Bahnnutzer* mit den „*wahlfreien*“ *Pkw-Nutzern*. Die „*nicht wahlfreien*“

⁴¹⁷ Für das Konstrukt der *Informationen* wurde in allen Modellen ein Parameter mit einem niedrigeren Wert als für den Parameter des Konstrukts der *Verbindungsqualität* bzw. des *Komforts* geschätzt.

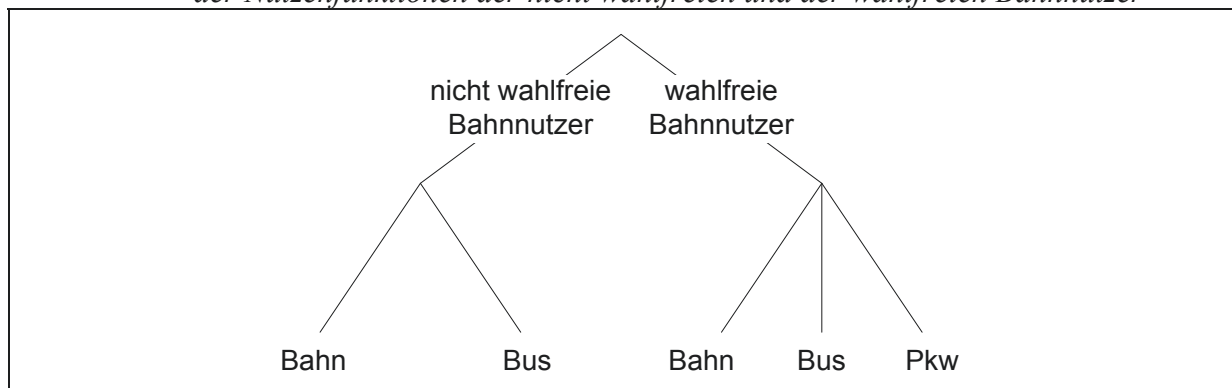
⁴¹⁸ Vgl. *Scarpa et al.* (2012), S. 1–2.

Pkw-Nutzer, d. h. die *Pkw-Nutzer*, die nicht zwischen den Alternativen gewechselt haben, werden aus allen folgenden Analysen ausgeschlossen.

7.6.1 Nicht wahlfreie Bahnnutzer und wahlfreie Bahnnutzer

Für die erste Stufe des Testverfahrens wurde ein zweistufiges Nested Logit-Modell mit den Daten der *nicht wahlfreien* und der *wahlfreien Bahnnutzer* geschätzt, dessen Aufbau in Abb. 7.7 veranschaulicht ist. In diesem Modell wird auf Zweigebene zwischen den beiden Gruppen unterschieden. Unterhalb des Zweigs der *nicht wahlfreien Bahnnutzer* befinden sich die beiden ÖV-Wahlalternativen und unterhalb des Zweigs der *wahlfreien Bahnnutzer* die beiden ÖV-Wahlalternativen und die Pkw-Alternative. Somit beinhaltet das Nested Logit-Modell insgesamt fünf elementare Alternativen.

Abb. 7.7: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer



Quelle: Eigene Darstellung.

Der Inklusivwertparameter (*IV*-Parameter) für den Zweig der *nicht wahlfreien Bahnnutzer* wurde auf 1 gesetzt. Die Schätzung des *IV*-Parameters für den Zweig der *wahlfreien Bahnnutzer* erfolgte relativ. Aufgrund der Gleichheitsrestriktion der Nutzenparameter erhöhte sich in diesem Modell die Anzahl der geschätzten Parameter im Vergleich zum MNL der *wahlfreien Bahnnutzer* (Tab. 7.8, Tab. 7.11 bzw. Tab. 7.14) lediglich um den zusätzlichen *IV*-Parameter.

Die Schätzung dieser zweistufigen NL-Modelle erfolgte getrennt nach den drei Subexperimenten *Verbindungsqualität*, *Komfort* und *Informationen* und resultierte für diese drei Modelle in Log-Likelihood-Werten von

-1553,64, -1579,34 bzw. -1234,14 bei 15, 15 bzw. 19 geschätzten Parametern.

Differenziert nach Subexperimenten wird in der ersten Stufe des Likelihood-Ratio-Tests das zweistufige NL-Modell mit den zugehörigen einzelnen MNL-Modellen (Tab. 7.8, Tab. 7.11 bzw. Tab. 7.14) anhand der Log-Likelihood-Werte verglichen. Die Ergebnisse der ersten Stufe dieses Tests sind in Tab. 7.17 dargestellt. Die Teststatistik λ_A überschreitet bei allen drei Subexperimenten den kritischen Wert, sodass die erste Subhypothese durchgehend abgelehnt und die zweite Stufe des Testverfahrens nicht mehr durchgeführt wurde. Aus diesem Grund wurde ebenfalls die Gesamthypothese, dass die Nutzenfunktionen zwischen den beiden Gruppen der *Bahnnutzer* gleich sind, abgelehnt. Wegen der unterschiedlichen Nutzenfunktionen erfolgt die weitere Analyse für diese beiden Nutzergruppen getrennt.

Tab. 7.17: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten

| Subexperiment | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ^2_{krit} ($\alpha=0,05$) |
|---------------------|---|---------------------------|-----------------------------------|
| Verbindungsqualität | $\lambda_A = -2[-1553,64 - (-654,25 - 880,09)] = 38,58$ | 11+14-15=10 | 18,31 |
| Komfort | $\lambda_A = -2[-1579,34 - (-707,23 - 851,69)] = 40,83$ | 11+14-15=10 | 18,31 |
| Informationen | $\lambda_A = -2[-1234,14 - (-610,74 - 608,43)] = 29,93$ | 15+18-19=14 | 23,68 |

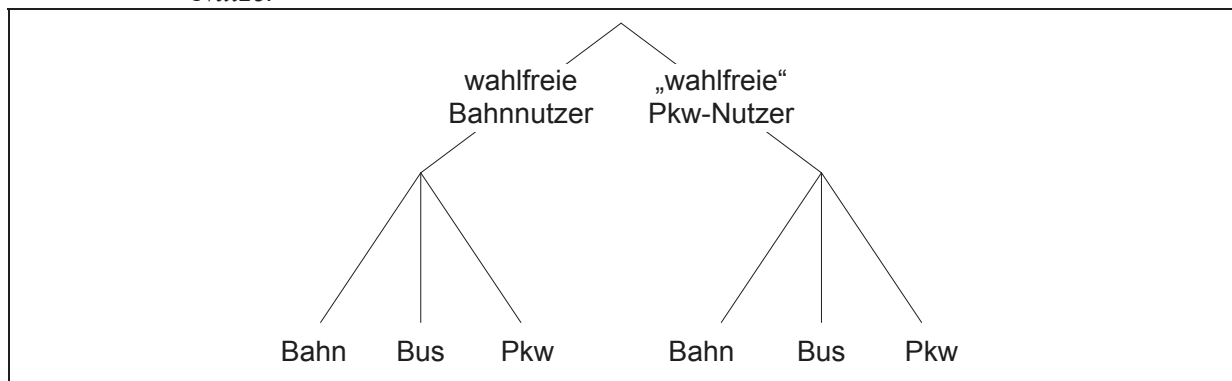
Quelle: Eigene Berechnungen.

7.6.2 Wahlfreie Bahnnutzer und „wahlfreie“ Pkw-Nutzer

Mit den Daten der *wahlfreien Bahnnutzer* und der „wahlfreien“ *Pkw-Nutzer* wurde für die erste Stufe des Testverfahrens ein gemeinsames zweistufiges Nested Logit-Modell geschätzt, dessen Aufbau in Abb. 7.8 dargestellt ist. Auf der Ebene der Zweige wird zwischen den beiden Gruppen unterschieden. Unterhalb der Zweige befinden sich jeweils alle drei Wahlalternativen, sodass insgesamt sechs elementare Alternativen existieren.

Der *IV*-Parameter für den Zweig der *wahlfreien Bahnnutzer* wurde auf 1 gesetzt, der *IV*-Parameter für den Zweig der „wahlfreien“ *Pkw-Nutzer* wurde relativ geschätzt. Die Anzahl der zu schätzenden Parameter erhöhte sich nur um diesen *IV*-Parameter.

Abb. 7.8: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer



Quelle: Eigene Darstellung.

Die drei Modelle der Subexperimente *Verbindungsqualität*, *Komfort* und *Informationen* mit 15, 15 bzw. 19 geschätzten Parametern resultierten in Log-Likelihood-Werten von -2196,15, -2183,37 bzw. -1971,58.

Die Ergebnisse der ersten Stufe des Testverfahrens des Likelihood-Ratio-Tests sind in Tab. 7.18 dargestellt. Bei allen drei Subexperimenten überschreitet die Teststatistik λ_A den kritischen Wert, sodass die erste Subhypothese und damit auch die Gesamthypothese abgelehnt wurden. Aufgrund der Ungleichheit der Nutzenfunktionen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* erfolgt die weitere Analyse getrennt nach Nutzergruppen.

Tab. 7.18: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer differenziert nach Subexperimenten

| Subexperiment | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ_{krit}^2 ($\alpha=0,05$) |
|---------------------|--|---------------------------|-----------------------------------|
| Verbindungsqualität | $\lambda_A = -2[-2196,15 - (-880,0 - 1234,22)]$ = 163,67 | 14+14-15=13 | 22,36 |
| Komfort | $\lambda_A = -2[-2183,37 - (-851,69 - 1263,01)]$ = 137,34 | 14+14-15=13 | 22,36 |
| Informationen | $\lambda_A = -2[-1971,58 - (-608,43 - 1278,12)]$ = 170,07 | 18+18-19=17 | 27,59 |

Quelle: Eigene Berechnungen.

7.7 Hausman-McFadden-Test der Unabhängigkeit von irrelevanten Alternativen

Mit dem Hausman-McFadden-Test der Unabhängigkeit von irrelevanten Alternativen (vgl. Kap. 3.5.1) wird ebendiese dem multinomialen Logit-Modell zugrunde liegende Annahme überprüft.

Im Folgenden wird der Test differenziert nach den Subexperimenten für die trimodalen Modelle, d. h. für die Modelle der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* durchgeführt. Bei diesem Test wird ein unrestringiertes Modell mit einem restringierten Modell verglichen. Das unrestringierte Modell entspricht in diesem Fall dem multinomialen Logit-Modell (Tab. 7.8, Tab. 7.9, Tab. 7.11, Tab. 7.12, Tab. 7.14 bzw. Tab. 7.15), in dem alle drei Alternativen vorhanden sind. Im restringierten Modell werden die Pkw-Alternative sowie alle Wahlsituationen, in denen diese Alternative gewählt wurde, eliminiert. Der Vergleich erfolgt anhand der Nutzenparameter (ohne Nutzenparameter der alternativenspezifischen Konstanten), Varianzen- und Kovarianzen. Es werden dabei nur Parameter berücksichtigt, die in beiden Modellen enthalten sind, d. h. die Pkw-spezifischen Parameter werden ausgeschlossen.

7.7.1 Wahlfreie Bahnnutzer

Das für das Modell der *Verbindungsqualität* errechnete Testergebnis beträgt $q = 29,96$. Die Anzahl der Freiheitsgrade entspricht der Anzahl der Parameter im restringierten Modell (ohne alternativenspezifische Konstanten), sodass ein kritischer Wert $\chi_{krit}^2 = 18,31$ (10 Freiheitsgrade) vorliegt, den das Testergebnis überschreitet.

Das Testergebnis für das Modell des *Komforts* beträgt $q = 45,11$. Bei einer gleichen Anzahl von Freiheitsgraden wie im vorherigen Modell liegt ein kritischer Wert $\chi_{krit}^2 = 18,31$ vor, welcher ebenfalls überschritten wird.

Für das Modell der *Informationen* wurde ein Testergebnis $q = 34,03$ ermittelt. Ein Vergleich mit dem kritischen Wert $\chi_{krit}^2 = 23,68$ bei 14 Freiheitsgraden zeigt, dass auch hierbei das Testergebnis den kritischen Wert überschreitet.

Aus der Überschreitung des kritischen Werts wurde eine Ablehnung der Annahme der Unabhängigkeit von irrelevanten Alternativen gefolgert. Da dem MNL-Modell diese Annahme zugrunde liegt, ist mit diesem Modell eine angemessene Modellierung der Daten nicht gegeben. Für die weitere Analyse erscheint daher ein Nested Logit-Modell geeigneter.

7.7.2 „Wahlfreie“ Pkw-Nutzer

Für das Modell der *Verbindungsqualität* liegt ein Testergebnis von $q = 13,78$ und für das Modell des *Komforts* von $q = 12,52$ vor. Bei jeweils 10 Freiheitsgraden beträgt der kritische Wert $\chi_{krit}^2 = 18,31$. Dieser Wert wird in beiden Fällen nicht überschritten.

Der Test für das Modell der *Informationen* führt zu einem Ergebnis von $q = 27,54$. Dieser Wert überschreitet den kritischen Wert $\chi_{krit}^2 = 23,68$ (14 Freiheitsgrade).

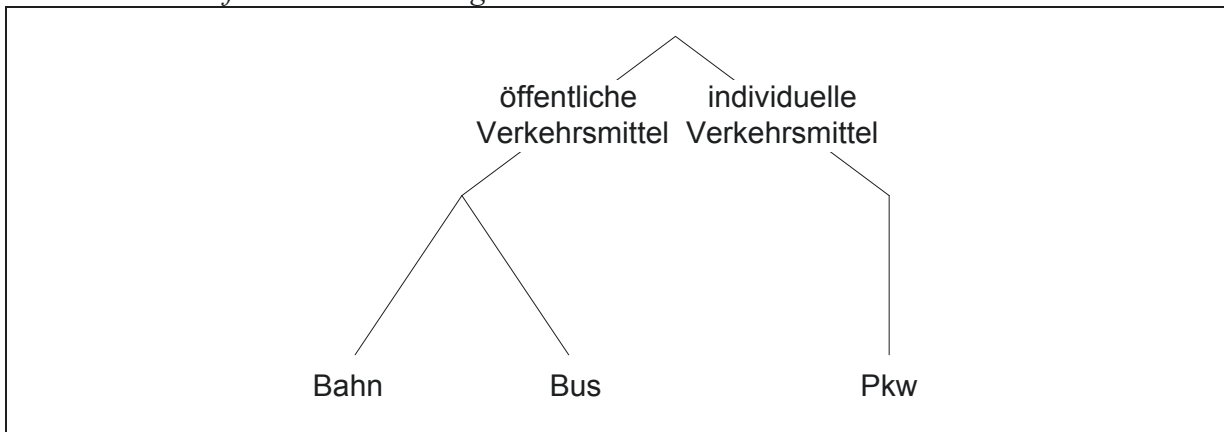
Da bei den Modellen der *Verbindungsqualität* und des *Komforts* die Annahme der Unabhängigkeit von irrelevanten Alternativen nicht abgelehnt wurde, erscheint eine Modellierung mit dem MNL-Modell grundsätzlich angemessen. Jedoch wurde ebendiese Annahme für das Modell der *Informationen* abgelehnt. Im Sinne einer gemeinsamen Modellgrundlage erscheint daher eine durchgehende Verwendung eines Nested Logit-Modells, d. h. auch für das Modell der *Verbindungsqualität* und des *Komforts*, angebracht.

7.8 Nested Logit-Modell

Aufgrund der Verletzung der IIA-Annahme wurden für die Gruppen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* Nested Logit-Modelle differenziert nach den drei unterschiedlichen Subexperimenten geschätzt. Der Aufbau dieser Modelle ist in Abb. 7.9 veranschaulicht. Wie der Abbildung zu entnehmen ist, wird in den Modellen auf der Ebene der Zweige zwischen öffentlichen und individuellen Verkehrsmitteln unterschieden.⁴¹⁹ Unterhalb des Zweigs der öffentlichen Verkehrsmittel befinden

⁴¹⁹ Grundsätzlich existieren drei Kombinationsmöglichkeiten zur Unterscheidung der Alternativen auf der Ebene der Zweige. Die hier beschriebene Unterscheidung zwischen

Abb. 7.9: Aufbau der Nested Logit-Modelle



Quelle: Eigene Darstellung.

sich die Bahn- und die Busalternative und unterhalb des Zweigs der individuellen Verkehrsmittel nur die Pkw-Alternative. Es handelt sich somit bei der Pkw-Alternative um eine degenerierte Alternative. Der *IV*-Parameter des individuellen Verkehrsmittels wurde auf 1 normalisiert.

7.8.1 Subexperiment der Verbindungsqualität

Die Schätzergebnisse der beiden Modelle der *Verbindungsqualität* sind in Tab. 7.19 dargestellt.⁴²⁰ Wie bereits bei den Schätzergebnissen der MNL-Modelle geben die *t*-Statistiken der Nutzenparameter β bzw. γ an, ob sich die Parameter signifikant von 0 unterscheiden. Abweichend davon gibt die *t*-Statistik des *IV*-Parameters an, ob eine Abweichung von 1 besteht.

Mithilfe des Likelihood-Ratio-Tests wurde überprüft, ob das NL-Modell durch den zusätzlichen Parameter eine signifikante Verbesserung im Vergleich zum MNL-Modell darstellt. Für die Gruppe der *wahlfreien Bahnnutzer* führt dieser Test zu $\lambda = -2[-880,09 - (-869,01)] = 22,04$ und für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* zu $\lambda = -2[-1234,22 - (-1232,09)] = 4,26$. Ein Vergleich mit dem kritischen Wert $\chi_{krit}^2 = 3,84$ (1 Freiheitsgrad) zeigt, dass das NL-Modell bei beiden Gruppen zu einer signifikanten

öffentlichen und individuellen Verkehrsmitteln führte jedoch durchgehend zu den besten Ergebnissen.

⁴²⁰ Der Vergleich des Nested Logit-Modells mit dem jeweiligen Konstantenmodell wurde für die beiden Nutzergruppen jeweils für alle Subexperimente durchgeführt und zeigte durchgehend eine deutliche Ablehnung des Konstantenmodells zugunsten des NL-Modells, d. h. die Gesamtsignifikanz der Modelle ist gegeben. Aufgrund der bereits deutlichen Ablehnung des Konstantenmodells zugunsten des MNL-Modells wird auf die Überprüfung anhand des Likelihood-Ratio-Tests nicht weiter eingegangen.

Verbesserung führt.⁴²¹

Bis auf wenige Ausnahmen konnten in beiden Modellen durchgehend signifikante Werte für die Nutzenparameter geschätzt werden. Bei den Ausnahmen im Modell der *wahlfreien Bahnnutzer* handelt es sich um Parameter, die sich auf die Pkw-Konstante sowie die zweite von insgesamt drei Ausprägungen der Attribute *Pünktlichkeit* und *Umsteigehäufigkeit* beziehen.

Tab. 7.19: *Nested Logit-Modell der Verbindungsqualität für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer*

| | | Wahlfreie Bahnnutzer | | „Wahlfreie“ Pkw-Nutzer | |
|--|--|-------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,391 | (4,90) | 0,215 | (3,15) |
| Pkw-Konstante | | 0,317 | (1,07) | 1,081 | (4,33) |
| Zeit ÖV | | -0,026 | (-6,74) | -0,063 | (-10,58) |
| Preis ÖV | | -0,169 | (-6,85) | -0,169 | (-2,37) |
| Zeit IV | | -0,032 | (-5,79) | -0,070 | (-10,30) |
| Kosten IV | | -0,126 | (-5,69) | -0,134 | (-3,28) |
| Komfort | | 0,054 | (3,49) | 0,072 | (3,53) |
| Informationen | | 0,037 | (2,56) | -0,002 | (-0,12) |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,146 | | 0,246 | |
| | 2. <i>3 bis 10 Minuten Verspätung</i> | 0,053 | (1,45) | 0,132 | (2,28) |
| | 3. <i>10 bis 20 Minuten Verspätung</i> | -0,199 | (-3,91) | -0,378 | (-5,42) |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | 0,180 | | 0,375 | |
| | 2. <i>1 x Umsteigen mit Anschlussgarantie</i> | -0,029 | (-0,76) | -0,028 | (-0,50) |
| | 3. <i>1 x Umsteigen ohne Anschlussgarantie</i> | -0,152 | (-3,16) | -0,347 | (-5,10) |
| Takt | 1. <i>alle 30 Minuten</i> | 0,309 | | 0,367 | |
| | 2. <i>alle 60 Minuten</i> | 0,086 | (2,17) | 0,086 | (1,53) |
| | 3. <i>alle 120 Minuten</i> | -0,395 | (-4,88) | -0,453 | (-6,06) |
| IV-Parameter der Verkehrsmittel | | | | | |
| Individualverkehr | (normalisiert) | 1,000 | | 1,000 | |
| Öffentlicher Verkehr | | 0,413 | (7,28) | 0,761 | (2,35) |
| Anzahl Wahlsituationen | | 1010 | | 1338 | |
| Anzahl Parameter | | 15 | | 15 | |
| Log-Likelihood-Wert des Modells | | -869,07 | | -1232,09 | |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | | -1052,81 | | -1432,40 | |

Quelle: *Eigene Berechnungen.*

⁴²¹ Für die Gruppe der „wahlfreien“ Pkw-Nutzer war diese Verbesserung aufgrund des Ergebnisses des Hausman-McFadden-Tests nicht zu erwarten.

Die nicht signifikanten Parameter im Modell der „*wahlfreien*“ *Pkw-Nutzer* beziehen sich auf das Konstrukt der *Informationen* und die zweite von insgesamt drei Ausprägungen der Attribute *Umsteigehäufigkeit* und *Takt*. Aus den nicht signifikanten Parametern der zweiten von insgesamt drei Attributsausprägungen wurde wie bislang gefolgert, dass sich der Nutzen linear mit den jeweiligen Ausprägungen dieser qualitativen Attribute verändert. Für das Konstrukt der *Informationen* im Modell der „*wahlfreien*“ *Pkw-Nutzer* konnte bereits beim MNL-Modell kein signifikanter Parameter geschätzt werden.

In beiden Modellen liegt der Wert des *IV*-Parameters im 0,1-Intervall, d. h. die Nested Logit-Modelle sind konsistent mit der Nutzenmaximierung (vgl. Kap. 3.3). Der *t*-Test zur Überprüfung der Annahme, dass der *IV*-Parameter den Wert 1 aufweist, führt im Modell der *wahlfreien Bahnnutzer* zu $t = (1-0,413)/\sqrt{(0,0065)} = 7,28$ und im Modell der „*wahlfreien*“ *Pkw-Nutzer* zu $t = (1-0,761)/\sqrt{(0,0103)} = 2,35$. Diese Werte sind größer als der kritische Wert $t_{krit} = 1,96$. Die *IV*-Parameter sind somit in beiden Modellen signifikant von 1 verschieden. Die Ergebnisse stehen im Einklang mit dem oben beschriebenen Likelihood-Ratio-Test zum Vergleich des NL-Modells mit dem MNL-Modell.

Die signifikanten Parameter weisen durchgehend plausible Vorzeichen auf. Im Modell der *wahlfreien Bahnnutzer* ist der Parameter des Konstrukts des *Komforts* wertmäßig größer als der Parameter des Konstrukts der *Informationen*. Die Bahnkonstante ist positiv, d. h. bei ansonsten gleichen Ausprägungen stiftet die Regionalbahn einen größeren Nutzen als der Regionalbus. Aus der Pkw-Konstante im Modell der „*wahlfreien*“ *Pkw-Nutzer* kann gefolgert werden, dass der Pkw einen größeren Nutzen als der Regionalbus mit durchschnittlichen qualitativen Attributen stiftet. Dahingegen kann aus dem nicht signifikanten Parameter der Pkw-Konstante im Modell der *wahlfreien Bahnnutzer* geschlossen werden, dass kein Unterschied zwischen dem Nutzen der Pkw-Alternative und dem Nutzen des Regionalbusses mit durchschnittlichen qualitativen Attributen besteht.

7.8.2 Subexperiment des Komforts

Die Schätzergebnisse der beiden Nested Logit-Modelle des *Komforts* sind in Tab. 7.20 dargestellt.

Der Likelihood-Ratio-Test zum Vergleich des NL-Modells mit dem MNL-Modell führt für die Gruppe der *wahlfreien Bahnnutzer* zu einem Ergebnis von $\lambda = -2[-851,69 - (-832,07)] = 39,23$ und für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* zu $\lambda = -2[-1263,01 - (-1262,63)] = 0,76$. Aus dem Vergleich mit dem kritischen Wert $\chi^2_{krit} = 3,84$ (1 Freiheitsgrad) folgt, dass das NL-Modell der *wahlfreien Bahnnutzer* zu einer signifikanten

Tab. 7.20: *Nested Logit-Modell des Komforts für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer*

| | | Wahlfreie Bahnnutzer | | „Wahlfreie“ Pkw-Nutzer | |
|---|--|-------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,258 | (4,78) | 0,376 | (4,34) |
| Pkw-Konstante | | 0,203 | (0,75) | 2,462 | (7,23) |
| Zeit ÖV | | -0,027 | (-6,36) | -0,078 | (-12,28) |
| Preis ÖV | | -0,164 | (-6,30) | -0,166 | (-2,40) |
| Zeit IV | | -0,029 | (-5,51) | -0,082 | (-11,54) |
| Kosten IV | | -0,124 | (-5,71) | -0,134 | (-3,51) |
| Verbindungsqualität Informationen | | 0,061 | (4,18) | 0,183 | (6,85) |
| Toilette im Zug/Bus | 1. <i>sauber</i> | 0,102 | | 0,197 | |
| | 2. <i>verschmutzt</i> | -0,075 | (-2,37) | -0,173 | (-2,72) |
| | 3. <i>nicht vorhanden</i> | -0,027 | (-0,90) | -0,024 | (-0,38) |
| Sauberkeit Zug/Bus innen | 1. <i>sauber</i> | 0,108 | | 0,317 | |
| | 2. <i>Boden klebrig, Papier liegt auf den Sitzen</i> | -0,108 | (-3,82) | -0,317 | (-6,42) |
| Sitzplatzverfügbarkeit im Zug/Bus | 1. <i>ganze Fahrt Sitzplatz</i> | 0,226 | | 0,357 | |
| | 2. <i>halbe Fahrt Sitzplatz</i> | 0,008 | (0,30) | 0,066 | (1,08) |
| | 3. <i>kein freier Sitzplatz</i> | -0,234 | (-4,68) | -0,424 | (-5,81) |
| Komfort der Sitzplätze | 1. <i>bequem</i> | 0,024 | | 0,218 | |
| | 2. <i>unbequem</i> | -0,024 | (-1,19) | -0,218 | (-4,56) |
| IV-Parameter der Verkehrsmittel | | | | | |
| Individualverkehr | (normalisiert) | 1,000 | | 1,000 | |
| Öffentlicher Verkehr | | 0,302 | (11,87) | 0,903 | (0,92) |
| Anzahl Wahlsituationen | | 945 | | 1422 | |
| Anzahl Parameter | | 15 | | 15 | |
| Log-Likelihood-Wert des Modells | | -832,07 | | -1262,63 | |
| Log-Likelihood-Wert des zugeh. Konstanten- modells | | -1002,54 | | -1505,61 | |

Quelle: *Eigene Berechnungen.*

Verbesserung führt, das NL-Modell der „*wahlfreien*“ *Pkw-Nutzer* hingegen nicht. Diese Testergebnisse waren aufgrund des Hausman-McFadden-Tests zu erwarten.

Für eine überwiegende Mehrzahl der Parameter konnten sowohl im Modell der *wahlfreien Bahnnutzer* als auch im Modell der „*wahlfreien*“ *Pkw-Nutzer* signifikante Werte geschätzt werden. Genau wie im ersten Subexperiment ist jedoch die Pkw-Konstante im Modell der *wahlfreien Bahnnutzer* nicht signifikant. Zudem ist im Modell der *wahlfreien Bahnnutzer* wie schon im zugehörigen MNL-Modell der Parameter, der sich auf den *Komfort der Sitzplätze* bezieht, nicht signifikant. In beiden NL-Modellen weist der Parameter, der sich auf die zweite von insgesamt drei Ausprägungen des Attributs *Sitzplatzverfügbarkeit* bezieht, keinen signifikanten Wert auf. Daher wurde erneut gefolgert, dass sich der Nutzen linear mit den Ausprägungen dieses Attributs verändert. Für den Parameter der zweiten Ausprägung des Attributs *Toilette im Zug/Bus* wurde ein signifikanter negativer Wert geschätzt, der Parameter der dritten Ausprägung ist hingegen nicht signifikant von 0 verschieden. Dies wurde dahingehend interpretiert, dass mit einer *nicht vorhandenen Toilette* ein höherer Nutzen als mit einer *verschmutzten Toilette* verbunden ist.

Die Vorzeichen der signifikanten Parameter sind plausibel. Der Parameter des Konstrukts der *Verbindungsqualität* ist wertmäßig größer als der Parameter des Konstrukts der *Informationen*. Die positive Bahnkonstante zeigt, dass bei ansonsten gleichen Ausprägungen die Regionalbahn einen größeren Nutzen als der Regionalbus stiftet. Bei der Gruppe der *wahlfreien Bahnnutzer* besteht kein Unterschied zwischen dem Nutzen des Pkw und dem Regionalbus mit durchschnittlichen qualitativen Ausprägungen; der Gruppe der „*wahlfreien*“ *Pkw-Nutzer* stiftet der Pkw hingegen einen größeren Nutzen als der Regionalbus mit durchschnittlichen qualitativen Attributen.

Der Wert des *IV*-Parameters liegt für die Modelle beider Gruppen im 0,1-Intervall. Der *t*-Test zur Überprüfung der Annahme, der *IV*-Parameter unterscheidet sich von 1, führt im Modell der *wahlfreien Bahnnutzer* zu $t = (1-0,302)/\sqrt{0,0035} = 11,87$ und im Modell der „*wahlfreien*“ *Pkw-Nutzer* zu $t = (1-0,903)/\sqrt{0,0112} = 0,92$. Der kritische Wert $t_{krit} = 1,96$ wird nur von dem ersten Ergebnis überschritten. Daraus folgt, dass sich der

IV-Parameter im Modell der *wahlfreien Bahnnutzer* signifikant von 1 unterscheidet, im Modell der „*wahlfreien*“ *Pkw-Nutzer* hingegen nicht. Diese Ergebnisse stehen im Einklang mit den Ergebnissen des Hausman-McFadden-Test in Kap. 7.7 und des Likelihood-Ratio-Tests zum Vergleich des NL-Modells mit dem MNL-Modell.

7.8.3 Subexperiment der Informationen

Die Ergebnisse der Nested Logit-Modelle der *Informationen* befinden sich in Tab. 7.21.

Tab. 7.21: *Nested Logit-Modell der Informationen für die wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer*

| | | | Wahlfreie Bahnnutzer | | „Wahlfreie“ Pkw-Nutzer | |
|--|----|--|----------------------------|-------------------------------|---------------------------|------------------------------|
| | | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | | 0,407 | (4,51) | 0,136 | (2,71) |
| Pkw-Konstante | | | 0,002 | (0,01) | 1,580 | (6,11) |
| Zeit ÖV | | | -0,036 | (-5,21) | -0,067 | (-10,40) |
| Preis ÖV | | | -0,242 | (-7,72) | -0,231 | (-3,14) |
| Zeit IV | | | -0,046 | (-5,65) | -0,077 | (-10,92) |
| Kosten IV | | | -0,126 | (-4,67) | -0,160 | (-3,87) |
| Verbindungsqualität | | | 0,073 | (3,69) | 0,141 | (6,97) |
| Komfort | | | 0,056 | (3,41) | 0,070 | (4,05) |
| Fahrplaninformationen | 1. | <i>vorhanden</i> | 0,040 | | 0,147 | |
| am Bahnhof/ an Bushaltestelle | 2. | nicht vorhanden | -0,040 | (-1,54) | -0,147 | (-4,17) |
| Informationen im Zug/Bus zu Grund und Dauer von Störungen | | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | 0,042 -0,002 0,005 | (1,65) (-0,07) (0,22) | 0,087 0,072 -0,041 | (2,66) (2,25) (-1,32) |
| Informationen am Bahnhof/an Bushalte- stelle zu Grund und Dauer von Störungen | | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | -0,019 0,029 0,007 | (-0,75) (1,10) (0,30) | 0,074 0,017 -0,038 | (2,28) (0,53) (-1,23) |
| Informationen im Zug/Bus zu aktuellen Anschlusszügen/-bussen | | Durchsageneffekt Anzeigeneffekt Interaktionseffekt | -0,027 -0,009 -0,011 | (-1,05) (-0,36) (-0,44) | -0,038 0,008 -0,014 | (-1,20) (0,26) (-0,44) |
| IV-Parameter der Verkehrsmittel | | | | | | |
| Individualverkehr | | (normalisiert) | 1,000 | | 1,000 | |
| Öffentlicher Verkehr | | | 0,334 | (8,81) | 0,586 | (5,58) |
| Anzahl Wahlsituationen | | | 760 | | 1374 | |
| Anzahl Parameter | | | 19 | | 19 | |
| Log-Likelihood-Wert des Modells | | | -595,34 | | -1268,95 | |
| Log-Likelihood-Wert des zugeh. Konstanten- modells | | | -742,47 | | -1496,52 | |

Quelle: *Eigene Berechnungen.*

Für die Gruppe der *wahlfreien Bahnnutzer* führt der Likelihood-Ratio-Test zum Vergleich des NL-Modells mit dem MNL-Modells zu $\lambda = -2[-608,43 - (-595,34)] = 26,18$ und für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* zu $\lambda = -2[-1278,12 - (-1268,95)] = 18,33$. Diese Werte sind größer als der kritische Wert $\chi_{krit}^2 = 3,84$ (1 Freiheitsgrad), d. h. das Nested Logit-Modell führt für beide Gruppen zu einer signifikanten Verbesserung. Diese Ergebnisse waren aufgrund des Hausman-McFadden-Tests zu erwarten.

Für die Modelle beider Gruppen konnten für die Zeit- und Preisattribute, die Konstrukte und die Bahnkonstante durchgehend signifikante Parameter geschätzt werden. Zudem weisen im Modell der „*wahlfreien*“ *Pkw-Nutzer* die Parameter der Pkw-Konstanten, der zweiten Ausprägung des Attributs *Fahrplaninformationen am Bahnsteig/an der Bushaltestelle*, der *Durchsagen-* und *Anzeigeneffekte* des Attributs *Informationen im Zug/Bus zu Grund und Dauer von Störungen* sowie des *Durchsageneffekts* des Attributs *Information am Bahnhof/an der Bushaltestelle zu Grund und Dauer von Störungen* signifikante Werte auf. Hinsichtlich der Signifikanz der Informationsattribute bestehen somit die gleichen Probleme wie in den MNL-Modellen in Kap. 7.5.3.

Der *IV*-Parameter liegt in beiden Modellen im 0,1-Intervall. Die Überprüfung mithilfe des *t*-Tests, ob der *IV*-Parameter signifikant von 1 abweicht, ergibt für die Gruppe der *wahlfreien Bahnnutzer* $t = (1 - 0,334) / \sqrt{0,0057} = 8,81$ sowie für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* $t = (1 - 0,586) / \sqrt{0,0055} = 5,58$. Die Testergebnisse überschreiten in beiden Fällen den kritischen Wert $t_{krit} = 1,96$. Somit liegt ein signifikant von 1 verschiedener *IV*-Parameter vor. Diese Ergebnisse stehen im Einklang mit denen des Hausman-McFadden-Tests in Kap. 7.7 und des Likelihood-Ratio-Tests zum Vergleich des NL-Modells mit dem MNL-Modell.

Die signifikanten Nutzenparameter weisen plausible Vorzeichen auf. Der Parameter des Konstrukts der *Verbindungsqualität* ist wertmäßig größer als der Parameter des Konstrukts des *Komforts*. Wie auch bei den anderen Modellen liegt eine positive Bahnkonstante vor, d. h. bei ansonsten gleichen Ausprägungen stiftet die Regionalbahn einen größeren Nutzen als der Regionalbus. Aus der nicht signifikanten Pkw-Konstante im

Modell der *wahlfreien Bahnnutzer* wurde gefolgert, dass kein Unterschied zwischen dem Nutzen der Pkw-Alternative und dem Regionalbus mit durchschnittlichen qualitativen Ausprägungen besteht. Dahingegen zeigt die positive Pkw-Konstante im Modell der „*wahlfreien*“ *Pkw-Nutzer*, dass die Pkw-Alternative einen größeren Nutzen als der Regionalbus mit durchschnittlichen qualitativen Ausprägungen stiftet.

7.9 Überprüfung der Gleichheit der Nutzenfunktionen der Nutzergruppen II

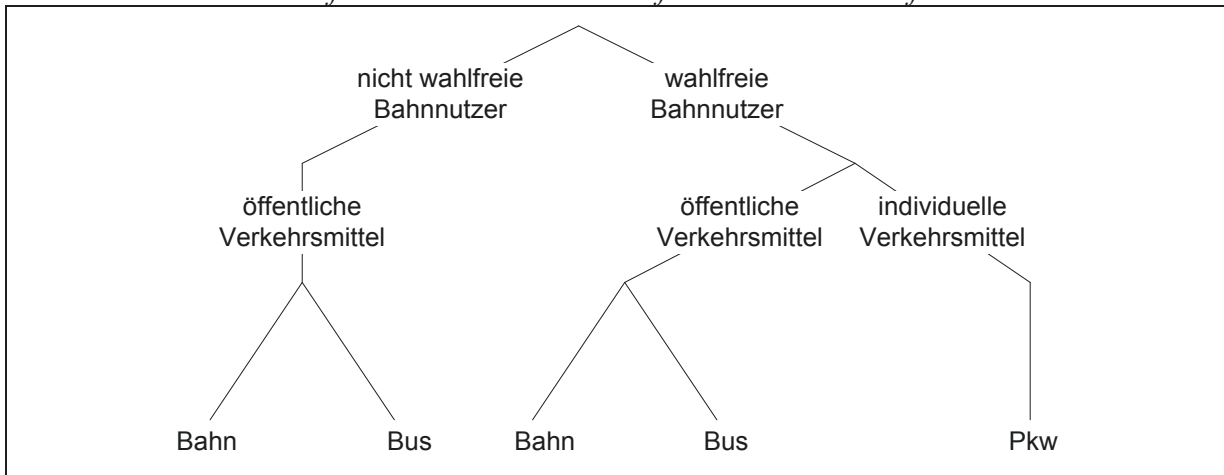
Aufgrund der Verwendung der Nested Logit-Modelle anstelle der MNL-Modelle ist eine erneute Überprüfung der Gleichheit der Nutzenfunktionen der Nutzergruppen erforderlich. Diese Überprüfung wird im Folgenden analog zu Kap. 7.6 differenziert nach Subexperimenten paarweise für die Nutzenfunktionen der *nicht wahlfreien* und der *wahlfreien Bahnnutzer* sowie der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* mithilfe eines zweistufigen Likelihood-Ratio-Tests durchgeführt.

7.9.1 Nicht wahlfreie und wahlfreie Bahnnutzer

Differenziert nach Subexperimenten wurde zur Überprüfung der ersten Subhypothese ein gemeinsames dreistufiges Nested Logit-Modell für die Gruppe der *nicht wahlfreien* und die Gruppe der *wahlfreien Bahnnutzer* geschätzt. Der Aufbau dieses Modells ist in Abb. 7.10 dargestellt.

Zwischen den beiden Nutzergruppen wird auf der Ebene der Äste unterschieden. Unterhalb des Asts der *nicht wahlfreien Bahnnutzer* befindet sich ein degenerierter Zweig, unterhalb dessen sich die beiden ÖV-Alternativen befinden. Ausgehend vom Ast der *wahlfreien Bahnnutzer* wird auf der Ebene der Zweige zwischen öffentlichen und individuellen Verkehrsmitteln unterschieden. Unterhalb der Zweigs der öffentlichen Verkehrsmittel befinden sich die Bahn- und die Busalternative und unterhalb des Zweigs der individuellen Verkehrsmittel die Pkw-Alternative als degenerierte Alternative. Wie im vergleichbaren Modell in Abb. 7.7 liegen somit insgesamt fünf elementare Alternativen vor.

Abb. 7.10: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer



Quelle: Eigene Darstellung.

Der *IV*-Parameter des individuellen Verkehrsmittels und des degenerierten Zweigs wurde auf 1 normalisiert. Zudem wurde der *IV*-Parameter des Asts für die Gruppe der *nicht wahlfreien Bahnnutzer* auf 1 gesetzt; die Schätzung des *IV*-Parameters für die andere Gruppe erfolgte relativ. Da die Nutzenparameter der beiden Gruppen auf Gleichheit restringiert wurden, umfasst das Modell im Vergleich zum einzelnen NL-Modell der *wahlfreien Bahnnutzer* (Tab. 7.19, Tab. 7.20 bzw. Tab. 7.21) einen zusätzlich zu schätzenden Parameter.

Die getrennt nach den Subexperimenten *Verbindungsqualität*, *Komfort* und *Informationen* geschätzten dreistufigen NL-Modelle umfassen 16, 16 bzw. 20 Parameter und weisen Log-Likelihood-Werte von -1533,71, -1547,07 bzw. -1216,76 auf.

In der ersten Stufe des Likelihood-Ratio-Tests wurde jeweils dieses dreistufige NL-Modell mit dem zugehörigen MNL-Modell der *nicht wahlfreien Bahnnutzer* (Tab. 7.8, Tab. 7.11 bzw. Tab. 7.14) und dem NL-Modell der *wahlfreien Bahnnutzer* (Tab. 7.19, Tab. 7.20 bzw. Tab. 7.21) anhand der Log-Likelihood-Werte verglichen. Die Ergebnisse der ersten Stufe dieses Tests befinden sich in Tab. 7.22.

Beim Subexperiment der *Verbindungsqualität* überschreitet die Teststatistik λ_A den kritischen Wert. Daraus folgt nicht nur eine Ablehnung der ersten Subhypothese, sondern auch eine Ablehnung der Gesamthypothese, d. h. es bestehen signifikante Unterschiede zwischen den Nutzen-

Tab. 7.22: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten

| Subexperiment | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ_{krit}^2 ($\alpha = 0,05$) |
|---------------------|---|---------------------------|-------------------------------------|
| Verbindungsqualität | $\lambda_A = -2[-1533,71 - (-654,23 - 869,07)] = 20,78$ | 11+15-16 = 10 | 18,31 |
| Komfort | $\lambda_A = -2[-1547,07 - (-707,23 - 832,07)] = 15,53$ | 11+15-16 = 10 | 18,31 |
| Informationen | $\lambda_A = -2[-1216,76 - (610,74 - 595,34)] = 21,34$ | 15+19-20 = 14 | 23,68 |

Quelle: Eigene Berechnungen.

funktionen der *nicht wahlfreien* und der *wahlfreien Bahnnutzer*. Bei diesem Subexperiment erschien daher eine getrennte Analyse erforderlich. Die Testergebnisse für die anderen beiden Subexperimente überschreiten die kritischen Werte hingegen nicht. Somit wird die Subhypothese, dass Gleichheit der Nutzenparameter bei möglichen Unterschieden zwischen den Skalierungsparametern besteht, nicht abgelehnt.

Für die zweite Stufe des Tests wurden differenziert nach den anderen beiden Subexperimenten gemeinsame zweistufige NL-Modelle geschätzt. Dieses Modell entspricht von der Struktur dem NL-Modell der *wahlfreien Bahnnutzer* in Abb. 7.9, d. h. auf der Ebene der Zweige wird zwischen öffentlichen und individuellen Verkehrsmitteln unterschieden.

Der *IV*-Parameter des Zweigs der individuellen Verkehrsmittel wurde auf 1 gesetzt; die Schätzung des *IV*-Parameters des anderen Zweigs erfolgte relativ. Das Modell des *Komforts* umfasst 15 und das Modell der *Informationen* 19 Parameter. Die Log-Likelihood-Werte betragen -1549,38 bzw. -1220,71.

Die Ergebnisse der zweiten Stufe des Likelihood-Ratio-Tests sind in Tab. 7.23 dargestellt und zeigen, dass die Teststatistik λ_B in beiden Fällen den kritischen Wert überschreitet. Somit wurde in beiden Fällen die zweite Subhypothese, dass gleiche Skalierungsparameter vorliegen, abgelehnt. Eine gemeinsame Analyse der beiden Gruppen wäre also für diese Subexperimente grundsätzlich möglich, wenn die Unterschiede zwischen den Skalierungsparametern berücksichtigt werden.

Tab. 7.23: Zweite Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien und der wahlfreien Bahnnutzer differenziert nach Subexperimenten

| Subexperiment | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ_{krit}^2 ($\alpha = 0,05$) |
|---------------|--|---------------------------|-------------------------------------|
| Komfort | $\lambda_B = -2[-1549,38 - (-1547,07)] = 4,63$ | 16-15 = 1 | 3,84 |
| Informationen | $\lambda_B = -2[-1220,71 - (-1216,76)] = 7,90$ | 20-19 = 1 | 3,84 |

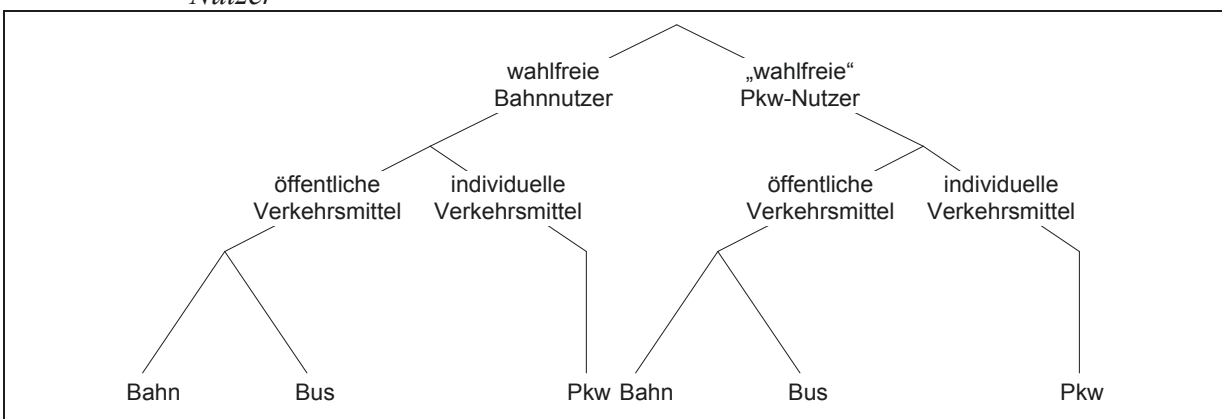
Quelle: Eigene Berechnungen.

7.9.2 Wahlfreie Bahnnutzer und „wahlfreie“ Pkw-Nutzer

Der Aufbau des dreistufigen NL-Modells zur Überprüfung der ersten Subhypothese für die Gruppen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* ist Abb. 7.11 veranschaulicht. Dieses Modell unterscheidet sich von dem entsprechenden Modell der beiden Gruppen der Bahnnutzer dahingehend, dass bei beiden Gruppen auf der Ebene der Zweige jeweils zwischen öffentlichen und individuellen Verkehrsmitteln unterschieden wird und insgesamt sechs elementare Alternativen vorliegen.

Neben den *IV*-Parametern der individuellen Verkehrsmittel wurde der *IV*-Parameter des Asts der Gruppe der „*wahlfreien*“ *Pkw-Nutzer* auf 1 gesetzt; die Schätzung des *IV*-Parameters für die Gruppe der *wahlfreien Bahnnutzer* erfolgte relativ.

Abb. 7.11: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer



Quelle: Eigene Darstellung.

Die Modelle der *Verbindungsqualität*, des *Komforts* und der *Informationen* weisen 17, 17 bzw. 21 Parameter auf; die Log-Likelihood-Werte betragen -2171,80, -2175,89 bzw. -1937,02.

Dieses dreistufige NL-Modell wurde mit den zugehörigen einzelnen NL-Modellen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* (Tab. 7.19, Tab. 7.20 bzw. Tab. 7.21) anhand der Log-Likelihood-Werte in der ersten Stufe des Likelihood-Ratio-Tests verglichen. Die Ergebnisse dieses Tests sind in Tab. 7.24 dargestellt.

Bei allen drei Subexperimenten überschreitet die Teststatistik λ_A den kritischen Wert. Deshalb wurde sowohl die erste Subhypothese als auch die Gesamthypothese der Gleichheit der Nutzenfunktionen der beiden Nutzergruppen durchgehend abgelehnt. Aus diesem Grund erfolgt die weitere Analyse für diese Gruppen getrennt.

Tab. 7.24: Erste Stufe des Likelihood-Ratio-Tests zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer und der „wahlfreien“ Pkw-Nutzer differenziert nach Subexperimenten

| Subexperiment | Teststatistik des Likelihood-Ratio-Tests | Anzahl der Freiheitsgrade | χ_{krit}^2 ($\alpha = 0,05$) |
|---------------------|---|---------------------------|--|
| Verbindungsqualität | $\lambda_A = -2[-2171,80 - (-869,07 - 1232,09)] = 141,29$ | 15+15-17 = 13 | 22,36 |
| Komfort | $\lambda_A = -2[-2175,89 - (-832,07 - 1262,62)] = 162,36$ | 15+15-17 = 13 | 22,36 |
| Informationen | $\lambda_A = -2[-1937,02 - (-595,34 - 1268,95)] = 145,45$ | 19+19-21 = 17 | 27,58 |

Quelle: Eigene Berechnungen.

7.10 Überprüfung der Prozessgleichheit zwischen den Subexperimenten

Die bisherige Analyse bezieht sich auf die in Kap. 7.4 dargelegten ersten beiden Fragestellungen. Als Ergebnis konnte bislang festgestellt werden, dass das Nested Logit-Modell zur Modellierung der vorliegenden trimodalen Wahlentscheidungen bis auf eine Ausnahme aufgrund der Verletzung der Annahme von irrelevanten Alternativen geeigneter erscheint als das multinomiale Logit-Modell. Auch für diese Ausnahme wird im Folgenden im Sinne einer einheitlichen Modellgrundlage das NL-Modell verwendet.

Zudem wurden Unterschiede zwischen den Nutzenfunktionen der drei Nutzergruppen festgestellt, sodass eine getrennte Analyse erforderlich ist. Lediglich für die *wahlfreien Bahnnutzer* und „*wahlfreien*“ *Pkw-Nutzer* wäre bei den Subexperimenten des *Komforts* und der *Information* eine gemeinsame Analyse unter Berücksichtigung von Unterschieden zwischen den Skalierungsparameter möglich. Aufgrund der bestehenden Unterschiede beim Subexperiment der *Verbindungsqualität* erfolgt jedoch auch für diese beiden Nutzergruppen eine durchgehend getrennte Analyse.

Aufbauend auf diesen Ergebnissen wird im Folgenden analysiert, ob zwischen den Subexperimenten Prozessgleichheit besteht. Diese liegt vor, wenn die drei Subexperimente das gleiche Wahlverhalten abbilden, d. h. wenn die Nutzenfunktionen einer Nutzergruppe sich zwischen den Subexperimenten nicht unterscheiden und es somit ohne Bedeutung ist, welches der Konstrukte in seine einzelnen Attribute aufgelöst wurde. Daher kann die Überprüfung der Prozessgleichheit als Überprüfung des hierarchischen Wahlmodells dieser konkreten Anwendung angesehen werden. Bei bestehender Prozessgleichheit können die Modelle aller drei Subexperimente zu einem gemeinsamen Modell mit allen Attributen und Konstrukten vereint werden, d. h. auf Basis der Daten der drei Subexperimente kann ein geeignetes gemeinsames Modell geschätzt werden.

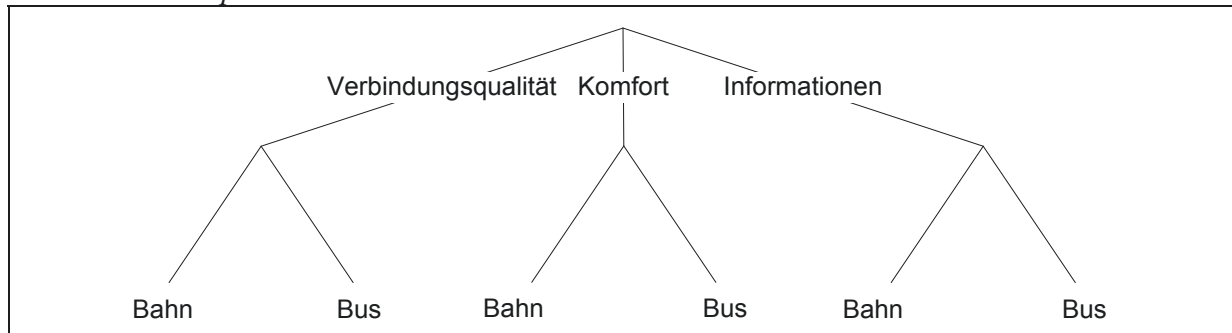
Die Überprüfung der Gleichheit der Nutzenfunktionen zwischen den Subexperimenten erfolgt wie in Kap. 3.5.2 und Kap. 4.3.3 beschrieben anhand eines zweistufigen Likelihood-Ratio-Tests differenziert nach Nutzergruppen.⁴²²

7.10.1 Nicht wahlfreie Bahnnutzer

Für die erste Stufe des Testverfahrens wurde für die drei Subexperimente ein gemeinsames zweistufiges NL-Modell geschätzt, in dem die Nutzenparameter, die in mehreren Subexperimenten verwendet werden, auf Gleichheit restringiert wurden. Der Aufbau dieses Modells ist in Abb. 7.12 dargestellt.

⁴²² Vgl. bezüglich der Überprüfung der Prozessgleichheit für die Gruppen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer*, bei der das MNL-Modell als Modellgrundlage verwendet wurde, Richter/Keuchel (2012b).

Abb. 7.12: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der nicht wahlfreien Bahnnutzer zwischen den Subexperimenten



Quelle: Eigene Darstellung.

In diesem Modell wird auf der Ebene der Zweige zwischen den drei Subexperimenten unterschieden. Unterhalb der Zweige befinden sich jeweils die Bahn- und die Busalternative, sodass insgesamt sechs elementare Alternativen vorliegen.

Der *IV*-Parameter des Zweigs der *Verbindungsqualität* wurde auf 1 gesetzt; die Schätzung der *IV*-Parameter der anderen beiden Subexperimente erfolgte relativ. Die Schätzergebnisse dieses Modells zeigt die Tab. 7.25. Das Modell weist insgesamt 30 Parameter und einen Log-Likelihood-Wert von -1977,90 auf.

Die Teststatistik zur Überprüfung der ersten Subhypothese beträgt $\lambda_A = -2[LL_{ges\ \mu} - (LL_{Verbindungsqualität} + LL_{Komfort} + LL_{Informationen})] = -2[-1977,90 - (-654,25 - 707,23 - 610,74)] = 11,33$. Dieser Wert überschreitet den kritischen Wert $\chi_{krit}^2 = 14,07$ ($11+11+15-30 = 7$ Freiheitsgrade) nicht. Die Subhypothese der Gleichheit der Nutzenparameter bei möglichen Unterschieden zwischen den Skalierungsparametern wurde daher nicht abgelehnt.

Für die zweite Stufe des Testverfahrens wurde für die drei Subexperimente ein gemeinsames MNL-Modell geschätzt, in dem die Nutzenparameter, die in mehreren Subexperimenten vorkommen, auf Gleichheit restringiert wurden und die Restriktion der Gleichheit der Skalierungsparameter zwischen den Subexperimenten durch die Verwendung des MNL-Modells gegeben ist.⁴²³ Die Ergebnisse dieses Modells

⁴²³ Die Gleichheit der Skalierungsparameter stellt eine der Restriktionen des multinomialen Logit-Modells dar. Das Modell entspricht einem zweistufigen NL-Modell, in dem die *IV*-Parameter aller Zweige auf 1 gesetzt werden.

Tab. 7.25: Gemeinsame Modelle mit allen Subexperimenten der nicht wahlfreien Bahnnutzer

| | | Zweistufiges NL-Modell | | MNL-Modell | |
|--|--|---------------------------|-----------|------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,538 | (10,57) | 0,598 | (14,95) |
| Zeit ÖV | | -0,060 | (-10,87) | -0,067 | (-16,39) |
| Preis ÖV | | -0,600 | (-6,56) | -0,671 | (-7,39) |
| Verbindungsqualität | | 0,176 | (8,26) | 0,202 | (11,84) |
| Komfort | | 0,154 | (7,95) | 0,173 | (9,89) |
| Informationen | | 0,072 | (4,62) | 0,079 | (4,99) |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,316 | | 0,326 | |
| | 2. 3 bis 10 Minuten Verspätung | 0,072 | (1,04) | 0,073 | (1,03) |
| | 3. 10 bis 20 Minuten Verspätung | -0,388 | (-5,63) | -0,399 | (-5,72) |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | 0,398 | | 0,409 | |
| | 2. 1 x Umsteigen mit Anschlussgarantie | -0,110 | (-1,58) | -0,110 | (-1,57) |
| | 3. 1 x Umsteigen ohne Anschlussgarantie | -0,288 | (-4,18) | -0,299 | (-4,29) |
| Takt | 1. <i>alle 30 Minuten</i> | 0,674 | | 0,700 | |
| | 2. alle 60 Minuten | 0,026 | (0,38) | 0,023 | (0,34) |
| | 3. alle 120 Minuten | -0,699 | (-9,64) | -0,724 | (-9,97) |
| Toilette im Zug/Bus | 1. <i>sauber</i> | 0,121 | | 0,138 | |
| | 2. verschmutzt | 0,022 | (0,38) | 0,026 | (0,40) |
| | 3. nicht vorhanden | -0,143 | (-2,49) | -0,164 | (-2,55) |
| Sauberkeit Zug/Bus innen | 1. <i>sauber</i> | 0,112 | | 0,128 | |
| | 2. Boden klebrig, Papier liegt auf den Sitzen | -0,112 | (-2,72) | -0,128 | (-2,79) |
| Sitzplatzverfügbarkeit im Zug/Bus | 1. <i>ganze Fahrt Sitzplatz</i> | 0,552 | | 0,634 | |
| | 2. halbe Fahrt Sitzplatz | 0,074 | (1,31) | 0,083 | (1,30) |
| | 3. kein freier Sitzplatz | -0,626 | (-7,29) | -0,717 | (-10,44) |
| Komfort der Sitzplätze | 1. <i>bequem</i> | 0,018 | | 0,020 | |
| | 2. unbequem | -0,018 | (-0,45) | -0,020 | (-0,44) |
| Fahrplaninformationen am Bahnhof/ an Bushaltestelle | 1. <i>vorhanden</i> | 0,096 | | 0,114 | |
| | 2. nicht vorhanden | -0,096 | (-2,32) | -0,114 | (-2,38) |
| Informationen im Zug/Bus zu Grund und Dauer von Störungen | Durchsageneffekt | 0,021 | (0,49) | 0,025 | (0,50) |
| | Anzeigeneffekt | 0,058 | (1,40) | 0,067 | (1,37) |
| | Interaktionseffekt | 0,003 | (0,06) | 0,003 | (0,05) |
| Informationen am Bahnhof/an Bushalte- stelle zu Grund und Dauer von Störungen | Durchsageneffekt | 0,045 | (1,08) | 0,053 | (1,09) |
| | Anzeigeneffekt | 0,115 | (2,70) | 0,134 | (2,74) |
| | Interaktionseffekt | -0,043 | (-1,03) | -0,051 | (-1,02) |
| Informationen im Zug/Bus zu aktuellen Anschlusszügen/-bussen | Durchsageneffekt | 0,015 | (0,35) | 0,019 | (0,38) |
| | Anzeigeneffekt | 0,031 | (0,75) | 0,034 | (0,70) |
| | Interaktionseffekt | -0,091 | (-2,21) | -0,108 | (-2,26) |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.25

| | Zweistufiges NL-Modell | | MNL-Modell |
|--|---------------------------|-----------|------------|
| | Par. | (t-Stat.) | |
| IV-Parameter der Subexperimente | | | |
| Verbindungsqualität (normalisiert) | 1,000 | | |
| Komfort | 0,867 | (1,39) | |
| Informationen | 0,833 | (1,86) | |
| Anzahl Wahlsituationen | 3685 | | 3685 |
| Anzahl Parameter | 30 | | 28 |
| Log-Likelihood-Wert des Modells | -1977,90 | | -1979,46 |
| Log-Likelihood-Wert des zugeh. Konstantenmodells | -2462,47 | | -2462,47 |

Quelle: Eigene Berechnungen.

sind ebenfalls in Tab. 7.25 dargestellt. Das Modell weist einen Log-Likelihood-Wert von -1979,46 auf und umfasst 28 Parameter.

Der Likelihood-Ratio-Test führt zu einem Ergebnis von $\lambda_B = -2[LL_{ges} - LL_{ges \mu}] = -2[-1979,46 - (-1977,90)] = 3,14$. Der kritische Wert $\chi^2_{krit} = 5,99$ ($30 - 28 = 2$ Freiheitsgrade) wird nicht überschritten, sodass sowohl die zweite Subhypothese als auch die Gesamthypothese nicht abgelehnt wurde, d. h. es liegt Gleichheit der Nutzenfunktionen vor.

Folglich ist die Prozessgleichheit für die drei Subexperimente der *nicht wahlfreien Bahnnutzer* gegeben, d. h. unabhängig vom verwendeten Subexperiment werden die gleichen Wahlentscheidungen abgebildet.

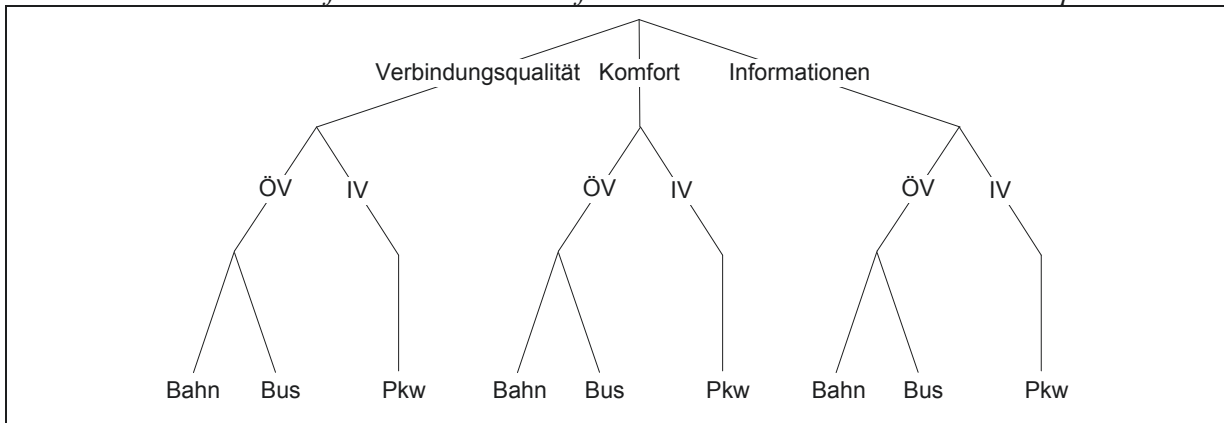
Eine weitere Überprüfung der Gleichheit der Skalierungsparameter der Subexperimente im dreistufigen NL-Modell ist mit einem *t*-Test möglich, bei dem getestet wird, ob sich der *IV*-Parameter des jeweiligen Subexperiments signifikant von 1 unterscheidet. Für den *IV*-Parameter des Subexperiments des *Komforts* führt dieser Test zu $t = (1 - 0,867) / \sqrt{0,0091} = 1,39$ und für den *IV*-Parameter der *Informationen* zu $t = (1 - 0,833) / \sqrt{0,0081} = 1,86$. Diese Werte überschreiten den kritischen Wert $t_{krit} = 1,96$ nicht, sodass die Hypothese der Gleichheit der Skalierungsparameter auch bei diesem Test nicht abgelehnt wurde.

Aus der bestehenden Prozessgleichheit folgt, dass die Modelle der drei Subexperimente dieser Nutzergruppe zu einem gemeinsamen Modell mit allen Attributen vereint werden können.

7.10.2 Wahlfreie Bahnnutzer

Für die erste Stufe des Testverfahrens wurde für die drei Subexperimente ein gemeinsames dreistufiges NL-Modell geschätzt, in dem die gemeinsamen Nutzenparameter der drei Subexperimente auf Gleichheit restringiert wurden. Der Aufbau dieses Modell ist in Abb. 7.13 veranschaulicht.

Abb. 7.13: Aufbau des künstlichen Nested Logit-Modells zur Überprüfung der Gleichheit der Nutzenfunktionen der wahlfreien Bahnnutzer zwischen den Subexperimenten



Quelle: Eigene Darstellung.

Die Unterscheidung zwischen den drei Subexperimenten erfolgt auf der Ebene der Äste und die Unterscheidung zwischen den öffentlichen und individuellen Verkehrsmitteln auf der Ebene der Zweige. Unterhalb der Zweige der öffentlichen Verkehrsmittel befinden sich jeweils die Bus- und die Bahnalternative und unterhalb der individuellen Verkehrsmittel jeweils die Pkw-Alternative als degenerierte Alternative. Insgesamt umfasst das Modell somit neun elementare Alternativen.

Die *IV*-Parameter des Asts der *Verbindungsqualität* und der Zweige der individuellen Verkehrsmittel wurden auf 1 gesetzt; die Schätzung der *IV*-Parameter der Äste der anderen beiden Subexperimente und der Zweige der öffentlichen Verkehrsmittel erfolgte relativ.

Darüber hinaus wurden den Nutzenfunktionen der beiden *ÖV*-Alternativen der Subexperimente der *Verbindungsqualität* und des *Komforts* die subexperimentspezifischen Konstanten K_1 und K_2 zugeordnet. Diese Dummyvariablen sollen mögliche Unterschiede zwischen dem durchschnittlichen Nutzen der qualitativen Attribute der Subexperimente berücksichtigen.

Im Modell der *nicht wahlfreien Bahnnutzer* waren derartige Konstanten nicht erforderlich, da eine Interpretation der Bahnkonstanten bei „ansonsten gleichen Ausprägungen“ möglich war. Dahingegen sind bei der Interpretation der Pkw-Konstante in den trimodalen Modellen ansonsten gleiche Ausprägungen nur dann gegeben, wenn alle Effektkodierungsvariablen den Wert 0 aufweisen, d. h. wenn die Pkw-Konstante relativ zum Nutzen der Referenzalternative mit durchschnittlichen qualitativen Attributen interpretiert wird (vgl. Kap. 7.5.1). Dieser Nutzen kann zwischen den Subexperimenten variieren, sodass beispielsweise der durchschnittliche Nutzen der qualitativen Attribute des *Komforts* höher ist als der durchschnittliche Nutzen der Attribute der *Verbindungsqualität*. Diese möglichen Unterschiede würden sich ansonsten in der bzw. den Pkw-Konstanten widerspiegeln. Die Verwendung von drei unterschiedlichen Pkw-Konstanten oder von einer gemeinsamen Pkw-Konstante und zwei subexperimentspezifischen Konstanten ist äquivalent. Letztere Variante wurde bevorzugt, da die möglichen Unterschiede durch das Design der Experimente bedingt sind.

Die Ergebnisse des gemeinsamen dreistufigen NL-Modells sind in Tab. 7.26 dargestellt. Das Modell umfasst 38 geschätzte Parameter und weist einen Log-Likelihood-Wert von -2300,37 auf.

Der Likelihood-Ratio-Test zur Überprüfung der ersten Subhypothese führt zu $\lambda_A = -2[LL_{ges\mu} - (LL_{Verbindungsqualität} + LL_{Komfort} + LL_{Informationen})] = -2[-2300,37 - (-869,07 - 832,07 - 595,34)] = 7,76$. Da dieser Wert den kritischen Wert $\chi_{krit}^2 = 19,68$ ($15 + 15 + 19 - 38 = 11$ Freiheitsgrade) nicht überschreitet, wurde die Subhypothese der Gleichheit der Nutzenparameter bei möglichen Unterschieden zwischen den Skalierungsparametern nicht abgelehnt.

Für die zweite Stufe des Likelihood-Ratio-Tests wurde für die drei Subexperimente ein gemeinsames zweistufiges NL-Modell geschätzt, in dem die gemeinsamen Nutzenparameter der Attribute und Konstrukte und die Skalierungsparameter⁴²⁴ auf Gleichheit restringiert wurden. Der Aufbau dieses Modells entspricht dem in Abb. 7.9.

⁴²⁴ Die Restriktion der Gleichheit der Skalierungsparameter zwischen den Subexperimenten ist durch das Auslassen der Ebene der Äste gegeben und entspricht einem dreistufigen NL-Modell, in dem die IV-Parameter aller Äste auf 1 gesetzt werden.

Tab. 7.26: Gemeinsame Modelle mit allen Subexperimenten der wahlfreien Bahnnutzer

| | | Dreistufiges NL-Modell | | Zweistufiges NL-Modell | |
|--|--|---------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,318 | (7,14) | 0,339 | (8,38) |
| Pkw-Konstante | | 0,165 | (0,87) | 0,028 | (0,14) |
| Zeit ÖV | | -0,027 | (-8,51) | -0,029 | (-10,83) |
| Preis ÖV | | -0,175 | (-8,87) | -0,185 | (-12,07) |
| Zeit IV | | -0,032 | (-7,77) | -0,033 | (-9,80) |
| Kosten IV | | -0,117 | (-7,49) | -0,127 | (-9,53) |
| Verbindungsqualität | | 0,061 | (5,66) | 0,068 | (6,27) |
| Komfort | | 0,046 | (4,96) | 0,049 | (5,20) |
| Informationen | | 0,038 | (4,16) | 0,038 | (4,46) |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,130 | | 0,126 | |
| | 2. 3 bis 10 Minuten Verspätung | 0,047 | (1,44) | 0,046 | (1,50) |
| | 3. 10 bis 20 Minuten Verspätung | -0,176 | (-4,30) | -0,172 | (-4,66) |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | 0,163 | | 0,161 | |
| | 2. 1 x Umsteigen mit Anschlussgarantie | -0,026 | (-0,76) | -0,025 | (-0,79) |
| | 3. 1 x Umsteigen ohne Anschlussgarantie | -0,138 | (-3,45) | -0,135 | (-3,72) |
| Takt | 1. <i>alle 30 Minuten</i> | 0,274 | | 0,270 | |
| | 2. alle 60 Minuten | 0,076 | (2,21) | 0,074 | (2,27) |
| | 3. alle 120 Minuten | -0,351 | (-5,88) | -0,344 | (-6,82) |
| Toilette im Zug/Bus | 1. <i>sauber</i> | 0,115 | | 0,120 | |
| | 2. verschmutzt | -0,084 | (-2,50) | -0,087 | (-2,59) |
| | 3. nicht vorhanden | -0,031 | (-0,93) | -0,033 | (-0,97) |
| Sauberkeit Zug/Bus innen | 1. <i>sauber</i> | 0,119 | | 0,125 | |
| | 2. Boden klebrig, Papier liegt auf den Sitzen | -0,119 | (-4,25) | -0,125 | (-4,64) |
| Sitzplatzverfügbarkeit im Zug/Bus | 1. <i>ganze Fahrt Sitzplatz</i> | 0,252 | | 0,262 | |
| | 2. halbe Fahrt Sitzplatz | 0,006 | (0,18) | 0,005 | (0,14) |
| | 3. kein freier Sitzplatz | -0,257 | (-5,35) | -0,267 | (-6,15) |
| Komfort der Sitzplätze | 1. <i>bequem</i> | 0,027 | | 0,028 | |
| | 2. unbequem | -0,027 | (-1,23) | -0,028 | (-1,26) |
| Fahrplaninformationen am Bahnhof/an Bushaltestelle | 1. <i>vorhanden</i> | 0,032 | | 0,039 | |
| | 2. nicht vorhanden | -0,032 | (-1,61) | -0,039 | (-1,60) |
| Informationen im Zug/Bus zu Grund und Dauer von Störungen | Durchsageneffekt | 0,033 | (1,69) | 0,039 | (1,64) |
| | Anzeigeneffekt | -0,001 | (-0,04) | -0,001 | (-0,02) |
| | Interaktionseffekt | 0,004 | (0,19) | 0,006 | (0,24) |
| Informationen am Bahnhof/an Bushalte- stelle zu Grund und Dauer von Störungen | Durchsageneffekt | -0,014 | (-0,73) | -0,018 | (-0,74) |
| | Anzeigeneffekt | 0,024 | (1,19) | 0,025 | (1,03) |
| | Interaktionseffekt | 0,005 | (0,28) | 0,007 | (0,29) |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.26

| | | Dreistufiges NL-Modell | | Zweistufiges NL-Modell | |
|---|--------------------|---------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Informationen im | Durchsageneffekt | -0,022 | (-1,14) | -0,026 | (-1,07) |
| Zug/Bus zu aktuellen | Anzeigeneffekt | -0,006 | (-0,34) | -0,006 | (-0,27) |
| Anschlusszügen/-bussen | Interaktionseffekt | -0,009 | (-0,48) | -0,011 | (-0,48) |
| Konstante K ₁ | | 0,095 | (0,52) | -0,069 | (-0,47) |
| Konstante K ₂ | | -0,081 | (-0,43) | -0,332 | (-2,26) |
| IV-Parameter der Verkehrsmittel | | | | | |
| Individualverkehr | (normalisiert) | 1,000 | | 1,000 | |
| Öffentlicher Verkehr | | | | 0,343 | (16,60) |
| ÖV Verbindungsqualität | | 0,369 | (11,15) | | |
| ÖV Komfort | | 0,337 | (12,41) | | |
| ÖV Informationen | | 0,263 | (18,10) | | |
| IV-Parameter der Subexperimente | | | | | |
| Verbindungsqualität | (normalisiert) | 1,000 | | | |
| Komfort | | 1,041 | (-0,27) | | |
| Informationen | | 0,764 | (2,14) | | |
| Anzahl Wahlsituationen | | 2715 | | 2715 | |
| Anzahl Parameter | | 38 | | 34 | |
| Log-Likelihood-Wert des Modells | | -2300,37 | | -2306,76 | |
| Log-Likelihood-Wert des zugeh. Konstanten- modells | | -2809,04 | | -2809,04 | |

Quelle: Eigene Berechnungen.

Auf der Ebene der Zweige wird zwischen den öffentlichen und individuellen Verkehrsmitteln unterschieden, unterhalb dessen sich die Bahn- und die Busalternative bzw. die Pkw-Alternative als degenerierte Alternative befinden. Insgesamt umfasst das Modell drei Alternativen. Für den Zweig der individuellen Verkehrsmittel wurde der *IV*-Parameter auf 1 gesetzt. Die Ergebnisse dieses Modells sind ebenfalls in Tab. 7.26 dargestellt. Für das Modell wurden 34 Parameter geschätzt und der Log-Likelihood-Wert beträgt -2306,76.

Der Likelihood-Ratio-Test zur Überprüfung der zweiten Subhypothese ergibt $\lambda_B = -2[LL_{ges} - LL_{ges \mu}] = -2[-2306,76 - (-2300,37)] = 12,77$. Der kritische Wert $\chi_{krit}^2 = 9,49$ (38-34 = 4 Freiheitsgrade) wird überschritten, wodurch die zweite Subhypothese abgelehnt wurde. Das bedeutet, dass Gleichheit zwischen den Nutzenparametern vorliegt, aber Unterschiede zwischen den Skalierungsparametern der Subexperimente bestehen.

Der *t*-Test zur Überprüfung der Gleichheit der Skalierungsparameter führt für den *IV*-Parameter des Subexperiments des *Komforts* zu $t = (1-1,041)/$

$\sqrt{0,0234} = -0,27$ und der *Informationen* zu $t = (1-0,764)/\sqrt{0,0121} = 2,14$. Der kritische Wert $t_{krit} = 1,96$ wird vom erstem Wert unterschritten und vom zweiten Wert überschritten. Demzufolge sind der Skalierungsparameter der *Verbindungsqualität* und der des *Komforts* gleich, wohingegen sich der Skalierungsparameter der *Informationen* von dem der *Verbindungsqualität* unterscheidet. In Übereinstimmung mit den Ergebnissen des Likelihood-Ratio-Tests wurde somit die Hypothese, dass alle Skalierungsparameter gleich sind, abgelehnt.

Ein gemeinsames Modell mit den Daten aller drei Subexperimente wäre für diese Nutzergruppe unter Berücksichtigung von Unterschieden zwischen den Skalierungsparametern möglich.

7.10.3 „Wahlfreie“ Pkw-Nutzer

Für die erste Stufe des Testverfahrens wurde für die drei Subexperimente ein gemeinsames dreistufiges NL-Modell geschätzt, welches hinsichtlich des Aufbaus, der Restriktionen, der Anzahl der Parameter und der Alternativen mit dem entsprechenden Modell der Gruppe der *wahlfreien Bahnnutzer* in Kap. 7.10.2 übereinstimmt.

Sämtliche Schätzergebnisse des dreistufigen NL-Modells sind in Tab. 7.27 dargestellt. Der Log-Likelihood-Wert beträgt -3771,04.

Das Ergebnis des Likelihood-Ratio-Tests zur Überprüfung der ersten Subhypothese beträgt $\lambda_A = -2[LL_{ges} - (LL_{Verbindungsqualität} + LL_{Komfort} + LL_{Informationen})] = -2[-3771,04 - (-1232,09 - 1262,63 - 1268,95)] = 14,74$. Der kritische Wert $\chi_{krit}^2 = 19,68$ (15+15+19-38 = 11 Freiheitsgrade) wird nicht überschritten, sodass die Subhypothese der Gleichheit der Nutzenparameter bei möglicherweise unterschiedlichen Skalierungsparametern nicht abgelehnt wurde.

Für die zweite Stufe des Likelihood-Ratio-Tests wurde für die drei Subexperimente ein gemeinsames zweistufiges NL-Modell geschätzt, das erneut hinsichtlich des Aufbaus, der Restriktionen, der Anzahl der Parameter und der Alternativen mit dem entsprechenden Modell der Gruppe der *wahlfreien Bahnnutzer* übereinstimmt. Die Schätzergebnisse sind ebenfalls in Tab. 7.27 enthalten. Das Modell weist einen Log-Likelihood-Wert von -3774,78 auf.

Tab. 7.27: Gemeinsame Modelle mit allen Subexperimenten der „wahlfreien“ Pkw-Nutzer

| | | Dreistufiges NL-Modell | | Zweistufiges NL-Modell | |
|---|--|---------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Bahnkonstante | | 0,199 | (5,39) | 0,223 | (5,87) |
| Pkw-Konstante | | 1,533 | (7,55) | 1,797 | (9,40) |
| Zeit ÖV | | -0,063 | (-11,23) | -0,069 | (-19,40) |
| Preis ÖV | | -0,168 | (-4,18) | -0,183 | (-4,38) |
| Zeit IV | | -0,070 | (-11,06) | -0,076 | (-19,05) |
| Kosten IV | | -0,127 | (-5,42) | -0,140 | (-5,97) |
| Verbindungsqualität | | 0,138 | (7,87) | 0,158 | (10,13) |
| Komfort | | 0,070 | (5,25) | 0,077 | (5,62) |
| Informationen | | 0,035 | (2,70) | 0,034 | (2,56) |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | 0,248 | | 0,245 | |
| | 2. 3 bis 10 Minuten Verspätung | 0,134 | (2,30) | 0,129 | (2,32) |
| | 3. 10 bis 20 Minuten Verspätung | -0,383 | (-5,47) | -0,374 | (-6,12) |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | 0,378 | | 0,372 | |
| | 2. 1 x Umsteigen mit Anschlussgarantie | -0,025 | (-0,44) | -0,025 | (-0,45) |
| | 3. 1 x Umsteigen ohne Anschlussgarantie | -0,353 | (-5,19) | -0,347 | (-5,70) |
| Takt | 1. <i>alle 30 Minuten</i> | 0,371 | | 0,365 | |
| | 2. alle 60 Minuten | 0,084 | (1,49) | 0,084 | (1,53) |
| | 3. alle 120 Minuten | -0,455 | (-6,12) | -0,449 | (-6,98) |
| Toilette im Zug/Bus | 1. <i>sauber</i> | 0,146 | | 0,158 | |
| | 2. verschmutzt | -0,137 | (-2,69) | -0,154 | (-2,80) |
| | 3. nicht vorhanden | -0,009 | (-0,18) | -0,004 | (-0,08) |
| Sauberkeit Zug/Bus innen | 1. <i>sauber</i> | 0,249 | | 0,277 | |
| | 2. Boden klebrig, Papier liegt auf den Sitzen | -0,249 | (-5,86) | -0,277 | (-6,81) |
| Sitzplatzverfügbarkeit im Zug/Bus | 1. <i>ganze Fahrt Sitzplatz</i> | 0,275 | | 0,305 | |
| | 2. halbe Fahrt Sitzplatz | 0,048 | (1,01) | 0,056 | (1,06) |
| | 3. kein freier Sitzplatz | -0,323 | (-5,32) | -0,361 | (-6,04) |
| Komfort der Sitzplätze | 1. <i>bequem</i> | 0,167 | | 0,184 | |
| | 2. unbequem | -0,167 | (-4,30) | -0,184 | (-4,60) |
| Fahrplaninformationen am Bahnhof/ an Bushaltestelle | 1. <i>vorhanden</i> | 0,143 | | 0,173 | |
| | 2. nicht vorhanden | -0,143 | (-4,22) | -0,173 | (-4,53) |
| Informationen im Zug/Bus zu Grund und Dauer von Störungen | Durchsageneffekt | 0,085 | (2,67) | 0,100 | (2,67) |
| | Anzeigeneffekt | 0,069 | (2,25) | 0,085 | (2,34) |
| | Interaktionseffekt | -0,043 | (-1,42) | -0,050 | (-1,37) |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.27

| | | Dreistufiges NL-Modell | | Zweistufiges NL-Modell | |
|--|--------------------|---------------------------|-----------|---------------------------|-----------|
| | | Par. | (t-Stat.) | Par. | (t-Stat.) |
| Informationen am Bahn- hof/an Bushaltestelle zu Grund und Dauer von Störungen | Durchsageneneffekt | 0,071 | (2,27) | 0,083 | (2,22) |
| | Anzeigeneffekt | 0,015 | (0,51) | 0,019 | (0,51) |
| | Interaktionseffekt | -0,037 | (-1,24) | -0,042 | (-1,17) |
| Informationen im Zug/ Bus zu aktuellen Anschlusszügen/-bussen | Durchsageneneffekt | -0,034 | (-1,12) | -0,041 | (-1,12) |
| | Anzeigeneffekt | 0,009 | (0,30) | 0,012 | (0,34) |
| | Interaktionseffekt | -0,015 | (-0,51) | -0,015 | (-0,41) |
| Konstante K_1 | | 0,277 | (1,66) | 0,493 | (3,15) |
| Konstante K_2 | | -0,185 | (-1,21) | -0,099 | (-0,65) |
| IV-Parameter der Verkehrsmittel | | | | | |
| Individualverkehr | (normalisiert) | 1,000 | | 1,000 | |
| Öffentlicher Verkehr | | | | 0,732 | (5,04) |
| ÖV Verbindungsqualität | | 0,774 | (2,26) | | |
| ÖV Komfort | | 0,677 | (3,87) | | |
| ÖV Informationen | | 0,576 | (6,07) | | |
| IV-Parameter der Subexperimente | | | | | |
| Verbindungsqualität | (normalisiert) | 1,000 | | | |
| Komfort | | 0,835 | (1,69) | | |
| Informationen | | 0,912 | (0,79) | | |
| Anzahl Wahlsituationen | | 4134 | | 4134 | |
| Anzahl Parameter | | 38 | | 34 | |
| Log-Likelihood-Wert des Modells | | -3771,04 | | -3774,78 | |
| Log-Likelihood-Wert des zugeh. Konstanten- modells | | -4442,54 | | -4442,54 | |

Quelle: Eigene Berechnungen.

Der Likelihood-Ratio-Test zur Überprüfung der zweiten Subhypothese führt zu $\lambda_B = -2[LL_{ges} - LL_{ges \mu}] = -2[-3774,78 - (-3771,04)] = 7,47$. Dieser Wert überschreitet den kritischen Wert $\chi_{krit}^2 = 9,49$ ($38 - 34 = 4$ Freiheitsgrade) nicht, sodass die zweite Subhypothese und die Gesamthypothese der Gleichheit der Nutzenfunktionen nicht abgelehnt wurde. Zwischen den Nutzenfunktionen der drei Subexperimente besteht also Prozessgleichheit, d. h. unabhängig vom verwendeten Subexperiment werden die gleichen Wahlentscheidungen abgebildet.

Die Ergebnisse des t -Tests zur Überprüfung der Gleichheit der Skalierungsparameter der Subexperimente im dreistufigen NL-Modell betragen für den IV-Parameter des Subexperiments des *Komforts* $t = (1 - 0,835) / \sqrt{0,0095} = 1,69$ und für den IV-Parameter der *Informationen* $t = (1 - 0,992) / \sqrt{0,0124} = 0,79$. Der kritische Wert $t_{krit} = 1,96$ wird von

keinem der Werte überschritten, sodass die Hypothese der Gleichheit der Skalierungsparameter auch bei diesem Test nicht abgelehnt wurde.

Aufgrund der bestehenden Prozessgleichheit können die Modelle der drei Subexperimente der „*wahlfreien*“ *Pkw-Nutzer* zu einem gemeinsamen Modell vereint werden.

7.11 Gemeinsames Modell

7.11.1 Nicht wahlfreie Bahnnutzer

Aufgrund der bestehenden Prozessgleichheit können die Modelle der *nicht wahlfreien Bahnnutzer* zu einem gemeinsamen Modell vereint werden. Dabei handelt es sich um das MNL-Modell in Tab. 7.25.

Die Gesamtsignifikanz dieses Modells wurde anhand eines Likelihood-Ratio-Tests überprüft. Dazu wurde der Log-Likelihood-Wert des MNL-Modells mit dem eines Konstantenmodells verglichen. Das entsprechende Testergebnis beträgt $\lambda = -2[-1979,46 - (-2462,47)] = 966,01$, welches den kritischen Wert $\chi^2_{krit} = 40,11$ ($28-1 = 27$ Freiheitsgrade) deutlich überschreitet. Die Gesamtsignifikanz des Modells ist damit gegeben.

In Bezug auf die Signifikanz der einzelnen Nutzenparameter bestehen keine Unterschiede zu den einzelnen MNL-Modellen (Tab. 7.8, Tab. 7.11 bzw. Tab. 7.14). Die nicht signifikanten Parameter beziehen sich überwiegend auf Parameter der zweiten von insgesamt drei Ausprägungen der qualitativen Attribute (*Pünktlichkeit*, *Umsteigehäufigkeit*, *Takt*, *Toilette im Zug/Bus* sowie *Sitzplatzverfügbarkeit*) sowie auf kodierte Effekte der *Informationen*. Aus den nicht signifikanten Parametern der mittleren Ausprägung der qualitativen Attribute wurde erneut auf eine lineare Veränderung des Nutzens aufgrund der Änderung der Ausprägungen dieser Attribute geschlossen.

Die Größenordnungen der Parameter der Konstrukte zeigen in Übereinstimmung mit den einzelnen Modellen, dass eine Veränderung der Ausprägung des Konstrukts der *Verbindungsqualität* zu der größten und eine Veränderung der Ausprägung des Konstrukts der *Informationen* zu der kleinsten Nutzenänderung führen. Für die Bahnkonstante wurde erneut ein

signifikanter positiver Parameter geschätzt, d. h. bei ansonsten gleichen Ausprägungen stiftet die Regionalbahn einen größeren Nutzen als der Regionalbus.

7.11.2 Wahlfreie Bahnnutzer

Da sich die Nutzenfunktionen der *wahlfreien Bahnnutzer* nur in Bezug auf die Skalierungsparameter unterscheiden, können die Modelle der drei Subexperimente unter Berücksichtigung dieser Unterschiede zu einem gemeinsamen Modell vereint werden. Dabei handelt es sich um das dreistufige Nested Logit-Modell in Tab. 7.26.

Der Likelihood-Ratio-Test zur Überprüfung der Gesamtsignifikanz führt zu $\lambda = -2[-2300,37 - (-2809,04)] = 1017,34$. Die Gesamtsignifikanz des Modells ist gegeben, da dieser Wert den kritischen Wert $\chi^2_{krit} = 51,00$ bei $38 - 2 = 36$ Freiheitsgraden nicht überschreitet.

Bezüglich der Signifikanz der einzelnen Nutzenparameter des gemeinsamen Modells besteht kein Unterschied zu den einzelnen NL-Modellen (Tab. 7.19, Tab. 7.20 und Tab. 7.21). Die nicht signifikanten Parameter beziehen sich überwiegend auf die zweite von insgesamt drei Ausprägungen einzelner qualitativer Attribute (*Pünktlichkeit*, *Umsteigehäufigkeit* und *Sitzplatzverfügbarkeit im Zug/Bus*) sowie auf die kodierten Effekte der *Informationen*. Aus den nicht signifikanten Parametern der jeweils mittleren Ausprägung wurde auf eine lineare Veränderung des Nutzens infolge einer Änderung der Ausprägung geschlossen. Ebenso besteht ein derartiger linearer Zusammenhang für die Ausprägungen des Attributs *Toilette im Zug/Bus*. Hierbei besteht jedoch der Unterschied darin, dass der Parameter der zweiten Ausprägung signifikant und negativ ist, der Parameter der dritten Ausprägung hingegen nicht signifikant ist. Dies bedeutet, dass eine *nicht vorhandene Toilette* einen größeren Nutzen als eine *verschmutzte Toilette* stiftet.

Die Ausprägungen der Parameter weisen die gleichen Größenordnungen wie in den einzelnen Modellen auf, d. h. mit einer Veränderung der Ausprägungen des Konstrukts der *Verbindungsqualität* ist die größte und mit einer Veränderung der Ausprägungen des Konstrukts der *Informationen* die kleinste Änderung des Nutzens verbunden.

Wie schon in den einzelnen NL-Modellen konnte auch im Gesamtmodell kein signifikanter Parameter für die Pkw-Konstante geschätzt werden. Daraus wurde gefolgert, dass zwischen dem Nutzen des Pkw und dem Nutzen des Regionalbusses mit durchschnittlichen qualitativen Attributen der *Informationen*⁴²⁵ kein signifikanter Unterschied besteht. Dahingegen wurde aus der signifikanten positiven Bahnkonstante geschlossen, dass die Regionalbahn bei sonst gleichen Ausprägungen einen größeren Nutzen als der Regionalbus stiftet.

Für die subexperimentspezifischen Konstanten K_1 und K_2 , die den beiden ÖV-Alternativen des Subexperiments der *Verbindungsqualität* bzw. des *Komforts* zugeordnet wurden, wurden keine signifikanten Parameter geschätzt. Daraus folgt, dass zwischen dem durchschnittlichen Nutzen der qualitativen Attribute der *Verbindungsqualität* bzw. des *Komforts* und der *Informationen* (Referenz-Subexperiment) kein signifikanter Unterschied besteht.

7.11.3 „Wahlfreie“ Pkw-Nutzer

Da für die Gruppe der „wahlfreien“ Pkw-Nutzer Prozessgleichheit zwischen den Subexperimenten besteht, können die Modelle der drei Subexperimente zu einem gemeinsamen Modell vereint werden. Bei diesem gemeinsamen Modell handelt es sich um das zweistufige NL-Modell in Tab. 7.27.

Eine Überprüfung der Gesamtsignifikanz des Modells erfolgte anhand eines Likelihood-Ratio-Tests, bei dem das Modell mit einem Konstantenmodell verglichen wurde. Das Testergebnis beträgt $\lambda = -2[-3774,78 - (-4442,54)] = 1335,53$. Der kritische Wert $\chi_{krit}^2 = 46,19$ (34-2 = 32 Freiheitsgrade) wird deutlich überschritten, sodass die Gesamtsignifikanz des Modells gegeben ist.

Ein Vergleich des Gesamtmodells mit den einzelnen NL-Modellen (Tab. 7.19, Tab. 7.20 bzw. Tab. 7.21) zeigt in Bezug auf die Signifikanz der einzelnen Nutzenparameter eine fast vollständige Übereinstimmung. Die nicht signifikanten Parameter beziehen sich überwiegend auf Effekt-

⁴²⁵ Das Subexperiment der *Informationen* stellt das Referenzsubexperiment dar, für das keine subexperimentspezifische Konstante geschätzt wurde.

kodierungsvariablen der zweiten von insgesamt drei Ausprägungen der qualitativen Attribute (*Umsteigehäufigkeit*, *Takt* und *Sitzplatzverfügbarkeit*) sowie auf einige kodierte Effekte der *Informationen*. Aus den nicht signifikanten Parametern der jeweils mittleren Ausprägung wurde erneut auf eine lineare Veränderung des Nutzens infolge der Änderung der Ausprägungen der Attribute geschlossen. Aus dem nicht signifikanten Parameter der dritten Ausprägungen des Attributs *Toilette im Zug/Bus* in Verbindung mit einem signifikanten negativen Parameter für die zweite Ausprägung wurde ebenfalls ein linearer Zusammenhang gefolgert, wobei jedoch eine *nicht vorhandene Toilette* einen größeren Nutzen als eine *verschmutzte Toilette* stiftet.

Aus den Werten der Parameter der Konstrukte geht hervor, dass eine Änderung der Ausprägung des Konstrukts der *Verbindungsqualität* mit der größten und eine Änderung der Ausprägung des Konstrukts der *Informationen* mit der geringsten Änderung des Nutzens verbunden ist. Diese Ergebnisse waren bereits in den einzelnen NL-Modellen dieser Nutzergruppe sowie auch in den Modellen der anderen Nutzergruppen zu beobachten.

Sowohl für die Bahnkonstante als auch für die Pkw-Konstante wurden signifikante positive Parameter geschätzt. Daraus wurde geschlossen, dass bei ansonsten gleichen Ausprägungen die Regionalbahn einen größeren Nutzen als der Regionalbus stiftet, sowie, dass der Pkw einen größeren Nutzen als ein Regionalbus mit durchschnittlichen qualitativen Informationsattributen (Referenz-Subexperiment) aufweist.

Für die subexperimentspezifische Konstante K_1 , welche den beiden ÖV-Alternativen des Subexperiments der *Verbindungsqualität* zugewiesen wurde, wurde ein signifikanter positiver Parameter geschätzt. Daraus wurde geschlossen, dass die durchschnittlichen qualitativen Attribute der *Verbindungsqualität* einen größeren Nutzen als die durchschnittlichen Attribute der *Informationen* (Referenz-Subexperiment) stiften. Für die subexperimentspezifische Konstante K_2 , die den beiden ÖV-Alternativen des Subexperiments des *Komforts* zugewiesen wurde, wurde hingegen kein signifikanter Parameter geschätzt, d. h. zwischen dem durchschnittlichen Nutzen der Attribute des *Komforts* und dem durchschnittlichen

Nutzen der qualitativen Attribute der *Informationen* (Referenz-Subexperiment) besteht kein signifikanter Unterschied.

7.12 Bestimmung von Zahlungsbereitschaften

Aufbauend auf den Modellergebnissen wurden nach (3.19) marginale Zahlungsbereitschaften für eine Veränderung der Ausprägungen der Qualitätsattribute und Konstrukte bestimmt. Dazu wurden die signifikanten Nutzenparameter⁴²⁶ der beiden öffentlichen Verkehrsmittel durch den Parameter des Preisattributs (*Preis ÖV*) und die signifikanten Nutzenparameter der Pkw-Alternative durch den Parameter des Kostenattributs (*Kosten IV*) dividiert. Die so ermittelten Zahlungsbereitschaften sind differenziert nach Nutzergruppen in Tab. 7.28 dargestellt.

Die Interpretation der Zahlungsbereitschaften erfolgt relativ zur Referenzalternative (Regionalbus mit durchschnittlichen qualitativen Eigenschaften). Ein negativer Wert bedeutet, dass die Befragten bereit sind, den entsprechenden Betrag für eine Veränderung der Qualität zu bezahlen (*willingness to pay*); ein positiver Wert bedeutet, dass eine Veränderung der Qualität eine Kompensation durch einen geringeren Preis erfordert (*willingness to accept*). Wie in Kap. 3.2 erörtert ist eine Veränderung der Qualität bei gleichzeitiger Veränderung des Preises um den Wert der entsprechenden Zahlungsbereitschaft nutzenneutral, sodass sich die Wahrscheinlichkeit der jeweiligen Alternative nicht verändert.

In Bezug auf die Zahlungsbereitschaften unterscheiden sich die drei Nutzergruppen dergestalt, dass die *nicht wahlfreien Bahnnutzer* die betragsmäßig kleinsten und die „*wahlfreien*“ *Pkw-Nutzer* die betragsmäßig größten Zahlungsbereitschaften aufweisen.

Bei einer Verkürzung der *Gesamtreisezeit* um eine Minute bei einer Fahrt mit einem öffentlichen Verkehrsmittel könnte der *Fahrpreis* der *nicht wahlfreien Bahnnutzer* um 0,10 €, der *wahlfreien Bahnnutzer* um 0,16 € und der „*wahlfreien*“ *Pkw-Nutzer* um 0,38 € angehoben werden. Umgekehrt müsste der *Fahrpreis* um den entsprechenden Betrag bei einer

⁴²⁶ Die Berechnung erfolgte ebenso für die Attribute, bei denen der Parameter der zweiten von insgesamt drei Ausprägungen nicht signifikant ist.

Tab. 7.28: Marginale Zahlungsbereitschaften differenziert nach Nutzergruppen

| | | Nicht wahlfreie Bahnnutzer | Wahlfreie Bahnnutzer | „Wahlfreie“ Pkw-Nutzer |
|--|--|----------------------------------|-------------------------|---------------------------|
| Bahnkonstante | | -0,89 € | -1,82 € | -1,22 € |
| Pkw-Konstante | | | n. s. | -12,88 € |
| Konstante K_1 | | | n. s. | -2,70 € |
| Konstante K_2 | | | n. s. | n. s. |
| Zeit ÖV (pro min) | | 0,10 € | 0,16 € | 0,38 € |
| Zeit IV (pro min) | | | 0,28 € | 0,55 € |
| Verbindungsqualität (pro Stufe) | | -0,30 € | -0,35 € | -0,86 € |
| Komfort (pro Stufe) | | -0,26 € | -0,26 € | -0,42 € |
| Informationen (pro Stufe) | | -0,12 € | -0,21 € | -0,19 € |
| Pünktlichkeit | 1. <i>pünktlich bzw. max. 3 Minuten Verspätung</i> | -0,49 € | -0,74 € | -1,34 € |
| | 2. 3 bis 10 Minuten Verspätung | (-0,11 €) | (-0,27 €) | -0,71 € |
| | 3. 10 bis 20 Minuten Verspätung | 0,59 € | 1,01 € | 2,05 € |
| Umsteigehäufigkeit | 1. <i>0 x Umsteigen</i> | -0,61 € | -0,93 € | -2,04 € € |
| | 2. 1 x Umsteigen mit Anschlussgarantie | (0,16 €) | (0,15 €) | (0,14 €) |
| | 3. 1 x Umsteigen ohne Anschlussgarantie | 0,45 € | 0,79 € | 1,90 € |
| Takt | 1. <i>alle 30 Minuten</i> | -1,04 € | -1,57 € | -2,00 € |
| | 2. alle 60 Minuten | (-0,03 €) | -0,44 € | (-0,46 €) |
| | 3. alle 120 Minuten | 1,08 € | 2,01 € | 2,46 |
| Toilette im Zug/Bus | 1. <i>sauber</i> | -0,21 € | -0,66 € | -0,87 € |
| | 2. verschmutzt | (-0,04 €) | 0,48 € | 0,84 € |
| | 3. nicht vorhanden | 0,24 € | (0,18 €) | (0,02 €) |
| Sauberkeit Zug/Bus innen | 1. <i>sauber</i> | -0,19 € | -0,68 € | -1,52 € |
| | 2. Boden klebrig, Papier liegt auf den Sitzen | 0,19 € | 0,68 € | 1,52 € |
| Sitzplatzverfügbarkeit im Zug/Bus | 1. <i>ganze Fahrt Sitzplatz</i> | -0,94 € | -1,44 € | -1,67 € |
| | 2. halbe Fahrt Sitzplatz | (-0,12 €) | (-0,03 €) | (-0,31 €) |
| | 3. kein freier Sitzplatz | 1,07 € | 1,47 € | 1,98 € |
| Komfort der Sitzplätze | 1. <i>bequem</i> | n. s. | n. s. | -1,01 € |
| | 2. unbequem | n. s. | n. s. | 1,01 € |
| Fahrplaninforma- tionen am Bahnhof/ an Bushaltestelle | 1. <i>vorhanden</i> | -0,17 € | n. s. | -0,95 € |
| | 2. nicht vorhanden | 0,17 € | n. s. | 0,95 € |
| Informationen im Zug/ Bus zu Grund und Dauer von Störungen | Durchsageneffekt | n. s. | n. s. | -0,55 € |
| | Anzeigeneffekt | n. s. | n. s. | -0,47 € |
| | Interaktionseffekt | n. s. | n. s. | n. s. |
| Informationen am Bahnhof/an Bushalte- stelle zu Grund und Dauer von Störungen | Durchsageneffekt | n. s. | n. s. | -0,45 € |
| | Anzeigeneffekt | n. s. | n. s. | n. s. |
| | Interaktionseffekt | n. s. | n. s. | n. s. |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.28

| | | Nicht wahlfreie Bahnnutzer | Wahlfreie Bahnnutzer | „Wahlfreie“ Pkw-Nutzer |
|----------------------------|--------------------|----------------------------------|-------------------------|---------------------------|
| Informationen im | Durchsagenereffekt | n. s. | n. s. | n. s. |
| Zug/Bus zu aktuellen | Anzeigeneffekt | n. s. | n. s. | n. s. |
| Anschlusszügen/ -bussen | Interaktionseffekt | 0,16 € | n. s. | n. s. |

Quelle: Eigene Berechnungen.

n. s. = nicht signifikant. Werte in Klammern beziehen sich auf nicht signifikante Parameter der zweiten von insgesamt drei Ausprägungen mit linearem Nutzenverlauf.

Verlängerung der *Gesamtreisezeit* um eine Minute gesenkt werden. Die Zahlungsbereitschaft für eine Verkürzung der Reisezeit bzw. die erforderliche Kompensation für deren Verlängerung wird auch als Wert der Zeit (*value of time* bzw. *value of travel time savings*) bezeichnet. Vielfach wird dazu der entsprechende Wert mit 60 multipliziert, um die Zahlungsbereitschaft für eine Veränderung von einer Stunde zu bestimmen. Zeitliche Veränderungen von einer Stunde liegen jedoch außerhalb der Veränderungen des Wahlexperiments, sodass eine (eher fragwürdige) Extrapolation erforderlich wäre.

Bei einer Fahrt mit dem Pkw liegt eine höhere Zahlungsbereitschaft für eine Verkürzung bzw. eine höhere erforderliche Kompensation für eine Verlängerung der *Gesamtreisezeit* vor. Sie beträgt 0,28 €/min für die Gruppe der *wahlfreien Bahnnutzer* und 0,55 €/min für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer*.

Eine Verbesserung der Konstrukte der *Verbindungsqualität*, des *Komfort* und der *Informationen* um eine entsprechende Anzahl von Stufen auf der Skala der zusammenfassenden Konstruktwerte ist den *nicht wahlfreien Bahnnutzern* 0,30 €/Stufe, 0,26 €/Stufe bzw. 0,12 €/Stufe, den *wahlfreien Bahnnutzern* 0,35 €/Stufe, 0,26 €/Stufe bzw. 0,21 €/Stufe und den „*wahlfreien*“ *Pkw-Nutzern* 0,86 €/Stufe, 0,42 €/Stufe bzw. 0,19 €/Stufe wert.

Als Indikatoren für die Wichtigkeit der einzelnen Attribute können die Spannweiten zwischen der Zahlungsbereitschaft für die unterste/schlechteste und der Zahlungsbereitschaft für die oberste/beste Ausprägung der qualitativen Attribute angesehen werden, die in Tab. 7.29 angegeben sind. Dabei ist zu beachten, dass die Spannweiten von dem Erhebungsdesign durch die Wahl der verwendeten Ausprägungen

abhängen.⁴²⁷ Da jedoch für die drei Nutzergruppen gleiche Ausprägungen der Attribute verwendet wurden, ist anhand der Zahlungsbereitschaften bzw. deren Spannweiten ein Vergleich möglich.

Tab. 7.29: Spannweiten der marginalen Zahlungsbereitschaften differenziert nach Nutzergruppen

| | Nicht wahlfreie Bahnutzer | Wahlfreie Bahnutzer | „Wahlfreie“ Pkw-Nutzer |
|--|---------------------------------|------------------------|---------------------------|
| Pünktlichkeit | 1,08 € | 1,75 € | 3,40 € |
| Umsteigehäufigkeit | 1,06 € | 1,72 € | 3,94 € |
| Takt | 2,12 € | 3,58 € | 4,46 € |
| Toilette im Zug/Bus | 0,45 € | 1,14 € | 1,71 € |
| Sauberkeit Zug/Bus innen | 0,38 € | 1,37 € | 3,03 € |
| Sitzplatzverfügbarkeit im Zug/Bus | 2,01 € | 2,91 € | 3,65 € |
| Komfort der Sitzplätze | n. s. | n. s. | 2,02 € |
| Fahrplaninformationen am Bahnhof/ an Bushaltestelle | 0,34 € | 0,36 € | 1,89 € |

Quelle: Eigene Berechnungen.

Die vergleichsweise hohe Zahlungsbereitschaft für die Veränderung der *Verbindungsqualität* spiegelt sich auch in den Zahlungsbereitschaften für die einzelnen Attribute der *Verbindungsqualität* wieder. Die drei Attribute *Pünktlichkeit*, *Umsteigehäufigkeit* und *Takt* gehören zu den vier wichtigsten qualitativen Attributen. Das Attribut *Takt* weist bei allen drei Nutzergruppen die höchste Spannweite der Zahlungsbereitschaft auf und reicht von 2,12 € bei den *nicht wahlfreien Bahnutzern* über 3,58 € bei den *wahlfreien Bahnutzern* zu 4,46 € bei den „*wahlfreien*“ *Pkw-Nutzern*. Hinsichtlich der Spannweite der Zahlungsbereitschaft ist die *Sitzplatzverfügbarkeit im Zug/Bus* bei den *nicht wahlfreien* und den *wahlfreien Bahnutzern* mit 2,01 € und 2,91 € das zweitwichtigste und bei den „*wahlfreien*“ *Pkw-Nutzern* mit 3,65 € das drittwichtigste qualitative Attribut. Für letztere Gruppe ist die *Umsteigehäufigkeit* mit einer Spannweite von 3,94 € das zweitwichtigste Attribut, welches für die *nicht wahlfreien* und *wahlfreien Bahnutzer* mit 1,06 € und 1,72 € an vierter Stelle steht. Die Spannweite der Zahlungsbereitschaft der *nicht wahlfreien Bahnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* für *vorhandene Fahrplaninformationen am Bahnsteig* ist mit 0,34 € und 1,90 € vergleichsweise niedrig.

⁴²⁷ Hensher/Rose/Greene (2005), S. 360 verweisen darauf, dass bereits die Zahlungsbereitschaften an sich durch die verwendeten Ausprägungen beeinflusst sein können.

Wie oben bereits geschrieben hängen die Spannweiten der Zahlungsbereitschaften von den Ausprägungen der Attribute ab. Für das Attribut *Takt* würde möglicherweise nicht mehr die größte Spannweite der Zahlungsbereitschaften bestehen, wenn die Unterschiede zwischen den Ausprägungen des Attributs geringer ausfielen, d. h. wenn anstelle eines *30- bis 120-Minuten-Takt* beispielsweise ein *45- bis 90-Minuten-Takt* verwendet würde.

Neben der nutzenneutralen Erhöhung eines Qualitätsattributs bei gleichzeitiger Anhebung des Fahrpreises ist es ebenso denkbar, die Erhöhung eines Qualitätsattributs nutzenneutral durch die Absenkung eines oder mehrerer anderer Qualitätsattribute zu kompensieren. Beispielsweise könnte – ausgehend vom durchschnittlichen Nutzen – für die Gruppe der *nicht wahlfreien Bahnnutzer* der zusätzliche Nutzen infolge eines *30-Minutentakts* der Regionalbahn (+0,701 NE (Nutzeneinheiten)) durch ein erforderliches *Umsteigen ohne Anschlussgarantie* (-0,299 NE) und eine um 6 Minuten verlängerte *Gesamtreisezeit* ($6 \text{ min} \times 0,067 \text{ NE/min} = 0,402 \text{ NE}$) kompensiert werden ($0,701 \text{ NE} - 0,299 \text{ NE} - 0,402 \text{ NE} = 0 \text{ NE}$). Da eine derartige Veränderung des Bahnangebots für die Gruppe der *nicht wahlfreien Bahnnutzer* nutzenneutral ist, würden sich die Wahlwahrscheinlichkeiten nicht ändern.

Dies bedeutet jedoch nicht, dass diese Veränderungen auch für die anderen beiden Gruppen nutzenneutral sind. Denn gleichzeitig würde der Nutzen der Bahnalternative für die Gruppe der *wahlfreien Bahnnutzer* durch die Verdichtung des *Takts* (+0,275 NE), das erforderliche *Umsteigen ohne Anschlussgarantie* (-0,138 NE) und die verlängerte *Gesamtreisezeit* ($6 \text{ min} \times -0,027 \text{ NE/min} = -0,162 \text{ NE}$) insgesamt sinken ($0,275 \text{ NE} - 0,138 \text{ NE} - 0,162 \text{ NE} = -0,025 \text{ NE}$), wodurch sich die Wahlwahrscheinlichkeiten veränderten. Ebenso würde für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* durch die Verdichtung des *Takts* (+0,365 NE), das erforderliche *Umsteigen ohne Anschlussgarantie* (-0,347 NE) und die verlängerte *Gesamtreisezeit* ($6 \text{ min} \times -0,069 \text{ NE/min} = -0,414 \text{ NE}$) der Nutzen der Bahnalternative insgesamt sinken ($0,365 \text{ NE} - 0,347 \text{ NE} - 0,414 \text{ NE} = -0,396 \text{ NE}$), sodass sich auch hier die Wahlwahrscheinlichkeiten veränderten. Derartige Veränderungen erfordern eine Simulation von Wahlwahrscheinlichkeiten.

7.13 Simulation von Wahlwahrscheinlichkeiten

Aufbauend auf den gemeinsamen Modellen (vgl. Kap. 7.11) wird im Folgenden differenziert nach Nutzergruppen eine exemplarische Simulation für mehrere Szenarien durchgeführt.

Da im gemeinsamen Modell der *nicht wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* keine Unterschiede zwischen den Skalierungsparametern der einzelnen Subexperimente berücksichtigt werden müssen, können die Ausprägungen der Attribute unterschiedlicher Konstrukte gleichzeitig variiert werden. Im gemeinsamen Modell der *wahlfreien Bahnnutzer* ist dies nicht möglich, da aufgrund der Unterschiede zwischen den Skalierungsparametern, die mit den Parametern der Attribute bzw. Konstrukte eines Subexperiments verknüpft sind, die Konstrukte nicht durch die zugehörigen Attribute ersetzt werden können und vice versa. Aus diesem Grund unterscheiden sich die Szenarien, die für die Gruppen der *nicht wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* entwickelt wurden, von denen der *wahlfreien Bahnnutzer*.

Eine Übersicht der sieben Szenarien, die für die Gruppen der *nicht wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* entwickelt wurden, ist in Tab. 7.30 dargestellt. In Abhängigkeit der Ausprägungen der Attribute und Konstrukte, die die Grundlage dieser Szenarien bilden, wurden individuen spezifische Wahlwahrscheinlichkeiten bestimmt. Die Mittelwerte dieser Wahlwahrscheinlichkeiten können als Wahlanteile interpretiert werden.⁴²⁸ Eine Übersicht der ermittelten Wahlanteile befindet sich in Tab. 7.31.

Grundsätzlich wurden die Gesamtreisezeit und der Fahrpreis der beiden ÖV-Alternativen von der aktuellen bzw. alternativen Fahrt mit der Bahn von den Befragungsdaten übernommen. Sofern für die alternative Fahrt mit der Bahn keine Daten vorlagen, wurden die entsprechenden Werte wie in Kap. 7.1 beschrieben aus den Werten der aktuellen Fahrt mit dem Pkw abgeleitet. Ebenso wurden, sofern vorhanden, die Konstruktbewertungen der aktuellen bzw. alternativen Fahrt als zusammenfassende Konstruktwerte übernommen. Bei einer fehlenden Konstruktbewertung wurde der

⁴²⁸ Vgl. Train (2009), S. 31 und Hensher/Rose/Greene (2005), S. 366.

Tab. 7.30: Szenarien der Simulation für die nicht wahlfreien Bahnnutzer und die „wahlfreien“ Pkw-Nutzer

| | Bahn | Bus |
|------------|--|---|
| Szenario 1 | <p><i>nicht wahlfreie Bahnnutzer:</i></p> <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> der aktuellen Fahrt mit der Bahn - zusammenfassende <u>Konstruktwerte</u>: Konstruktbewertungen der aktuellen Fahrt (sofern angegeben, ansonsten Mediane der Konstruktbewertung) <p><i>„wahlfreie“ Pkw-Nutzer:</i></p> <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> der alternativen Fahrt mit der Bahn (sofern angegeben, ansonsten wie in Kap. 7.1 beschrieben aus den Werten der aktuellen Fahrt mit dem Pkw abgeleitet) - zusammenfassende <u>Konstruktwerte</u>: Konstruktbewertungen der alternativen Fahrt mit der Bahn (sofern angegeben, ansonsten Mediane der Konstruktbewertung) | identische Ausprägungen wie Bahnalternative in Szenario 1 |
| Szenario 2 | identische Ausprägungen wie Bahnalternative in Szenario 1, aber <u>Erhöhung des Fahrpreises</u> um 1 % | identische Ausprägungen wie Busalternative in Szenario 1 |
| Szenario 3 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Erhöhung der Gesamtreisezeit</u> um 10 Minuten |
| Szenario 4 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Reduzierung des Fahrpreises</u> um 10 % |
| Szenario 5 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Erhöhung der Gesamtreisezeit</u> um 10 Minuten und <u>Reduzierung des Fahrpreises</u> um 10 % |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.30

| | Bahn | Bus |
|------------|--|---|
| Szenario 6 | <ul style="list-style-type: none"> - <u>Gesamtreisezeit und Fahrpreis</u> wie Bahnalternative in Szenario 1 - <u>Pünktlichkeit</u>: <i>pünktlich</i> bzw. <i>max. 3 Minuten Verspätung</i> - <u>Umsteigehäufigkeit</u>: <i>0 x Umsteigen</i>, sofern bei der beschriebenen (aktuellen bzw. alternativen) Fahrt mit der Bahn kein Umsteigen erforderlich ist, ansonsten <i>1 x Umsteigen ohne Anschlussgarantie</i> - <u>Takt</u>: <i>alle 60 Minuten</i> - <u>Toilette im Zug</u>: <i>verschmutzt</i> - <u>Sauberkeit Zug innen</u>: <i>sauber</i> - <u>Sitzplatzverfügbarkeit im Zug</u>: <i>ganze Fahrt Sitzplatz (50 %), halbe Fahrt Sitzplatz (25 %), kein freier Sitzplatz (25 %)</i> - <u>Komfort der Sitzplätze</u>: <i>bequem</i> - zusammenfassender <u>Konstruktwert Informationen</u>: wie in Szenario 1 | <ul style="list-style-type: none"> - <u>Gesamtreisezeit und Fahrpreis</u> wie Busalternative in Szenario 3 (<i>Erhöhung der Gesamtreisezeit um 10 Minuten</i>) - <u>Pünktlichkeit</u>: <i>3 bis 10 Minuten Verspätung</i> - <u>Umsteigehäufigkeit</u>: <i>0 x Umsteigen</i>, sofern bei der beschriebenen (aktuellen bzw. alternativen) Fahrt mit der Bahn kein Umsteigen erforderlich ist, ansonsten <i>0 x Umsteigen (50 %) bzw. 1 x Umsteigen ohne Anschlussgarantie (50 %)</i> - <u>Takt</u>: <i>alle 30 Minuten</i> - <u>Toilette im Bus</u>: <i>nicht vorhanden</i> - <u>Sauberkeit Bus innen</u>: <i>sauber</i> - <u>Sitzplatzverfügbarkeit im Bus</u>: <i>ganze Fahrt Sitzplatz</i> - <u>Komfort der Sitzplätze</u>: <i>bequem</i> - zusammenfassender <u>Konstruktwert Informationen</u>: wie in Szenario 1 |
| Szenario 7 | identische Ausprägungen wie Bahnalternative in Szenario 6, aber <u>Toilette im Zug</u> : <i>sauber</i> | identische Ausprägungen wie Busalternative in Szenario 6 |

Quelle: Eigene Darstellung.

Median der Konstruktbewertung der jeweiligen Nutzergruppe (Abb. 7.4 bis Abb. 7.6) verwendet. Die Gesamtreisezeit und die Fahrtkosten (inkl. sonstiger Kosten) der Pkw-Alternative wurden von der alternativen bzw. aktuellen Fahrt mit dem Pkw übernommen und in den Szenarien nicht variiert.

In den ersten fünf Szenarien wurden die qualitativen Attribute indirekt in Form der zusammenfassenden Konstruktwerte berücksichtigt. Im ersten Szenario wurde keine Unterscheidung hinsichtlich der Ausprägungen der beiden ÖV-Alternativen vorgenommen, sodass sich diese beiden Alternativen nur durch die alternativenspezifische Konstante unterscheiden. Im Ergebnis betragen bei der Gruppe der *nicht wahlfreien Bahnnutzer* die

Wahlanteile der Bahn und des Busses 64,5 % bzw. 35,5 %. Bei der Gruppe der „wahlfreien“ Pkw-Nutzer betragen die Wahlanteile der Bahn, des Busses und des Pkw 33,8 %, 24,9 % bzw. 41,3 %.

Im zweiten bis fünften Szenario wurden die Gesamtreisezeit und/oder der Fahrpreis jeweils eines öffentlichen Verkehrsmittels verändert, darüber hinaus unterscheiden sich diese Szenarien vom ersten Szenario nicht, sodass Änderungen der Wahlanteile im Vergleich dazu analysiert werden können.

Ein Vergleich der Wahlanteile des ersten und zweiten Szenarios zeigt, dass eine Erhöhung des Fahrpreises der Bahn um 1 % bei der Gruppe der *nicht wahlfreien Bahnnutzer* zu einem Rückgang des Wahlanteils dieser Alternative um 0,56 Prozentpunkte (63,96 % - 64,52 % = -0,56 %) führt; der Wahlanteil des Busses steigt um denselben Betrag. Dies entspricht einer Preiselastizität von $(-0,56 \%/64,52 \%) / 1 \% = -0,87$ (unelastische Nachfrage). Bei der Gruppe der „wahlfreien“ Pkw-Nutzer führt diese Preiserhöhung zu einem Rückgang des Wahlanteils der Bahn um 0,18 Prozentpunkte (33,60 %-33,78 % = -0,18 %); der Wahlanteil des Busses steigt um 0,10 Prozentpunkte und der Wahlanteil des Pkw um 0,08 Prozentpunkte. Somit beträgt die Preiselastizität $(-0,18 \%/33,78 \%) / 1 \% = -0,52$ (unelastische Nachfrage). Diese Preiselastizitäten erscheinen zwar relativ hoch, die Größenordnungen können dennoch als realistisch angesehen werden.⁴²⁹

Tab. 7.31: Mittelwert der Wahlwahrscheinlichkeiten der Alternativen differenziert nach den nicht wahlfreien Bahnnutzern und den „wahlfreien“ Pkw-Nutzern

| | Nicht wahlfreie Bahnnutzer | | „Wahlfreie“ Pkw-Nutzer | | |
|------------|----------------------------|--------|------------------------|--------|--------|
| | Bahn | Bus | Bahn | Bus | Pkw |
| Szenario 1 | 64,5 % | 35,5 % | 33,8 % | 24,9 % | 41,3 % |
| Szenario 2 | 64,0 % | 36,0 % | 33,6 % | 25,0 % | 41,4 % |
| Szenario 3 | 78,1 % | 21,9 % | 42,4 % | 12,1 % | 45,5 % |
| Szenario 4 | 58,7 % | 41,3 % | 32,7 % | 26,5 % | 40,7 % |
| Szenario 5 | 73,5 % | 26,5 % | 41,7 % | 13,1 % | 45,2 % |
| Szenario 6 | 72,1 % | 27,9 % | 18,1 % | 17,5 % | 64,4 % |
| Szenario 7 | 74,1 % | 25,9 % | 23,7 % | 15,3 % | 61,0 % |

Quelle: Eigene Berechnungen.

⁴²⁹ Vgl. Bastians (2009), S. 128–147 für eine Meta-Analyse von Elastizitäten im öffentlichen Verkehr. Des Weiteren ist zu beachten, dass im Modell kein (absoluter) Nachfragerückgang abgebildet wird, sondern eine Veränderung der Verkehrsmittelwahl.

Durch eine Erhöhung der Gesamtreisezeit des Busses um 10 Minuten (Szenario 3) sinkt im Vergleich zum ersten Szenario bei den *nicht wahlfreien Bahnnutzern* der Wahlanteil des Busses um 13,59 Prozentpunkte; der Wahlanteil der Bahn steigt um denselben Betrag. Bei den „*wahlfreien*“ *Pkw-Nutzern* führt diese Veränderung zu einer Senkung des Wahlanteils des Busses um 12,79 Prozentpunkte und einer Erhöhung der Wahlanteile der Bahn und des Pkw um 8,58 bzw. 4,21 Prozentpunkte.

Eine Senkung des Fahrpreises des Busses um 10 % (Szenario 4) führt bei den *nicht wahlfreien Bahnnutzern* zu einer Steigerung des Wahlanteils des Busses um 5,83 und bei den „*wahlfreien*“ *Pkw-Nutzern* um 1,63 Prozentpunkte. Bei letzterer Gruppe vermindern sich die Wahlanteile der Bahn und des Pkw um 1,04 bzw. 0,60 Prozentpunkte.

Bei einer Erhöhung der Gesamtreisezeit des Busses um 10 Minuten und einer gleichzeitigen Preissenkung um 10 % (Szenario 5) sinkt der Wahlanteil des Busses bei den *nicht wahlfreien Bahnnutzern* um 8,94 und bei den „*wahlfreien*“ *Pkw-Nutzern* um 11,79 Prozentpunkte. Die Wahlanteile der Bahn und des Pkw steigen bei den „*wahlfreien*“ *Pkw-Nutzern* um 7,90 bzw. 3,89 Prozentpunkte.

Abweichend von den ersten fünf Szenarien wurden im sechsten Szenario der Bahn- und der Busalternative unterschiedliche (hypothetische) Ausprägungen der qualitativen Attribute zugeordnet: In diesem Szenario ist die Regionalbahn pünktlicher, der Regionalbus fährt häufiger als die Bahn. Bezüglich der Umsteigehäufigkeit ermöglichen sowohl die Bahn als auch der Bus eine Direktverbindung, wenn bei der aktuellen bzw. alternativen Bahnfahrt kein Umsteigen erforderlich ist. Zudem ermöglicht der Regionalbus mehr Direktverbindungen dergestalt, dass (zufällig ausgewählte) 50 % der Befragten den Regionalbus ohne Umsteigen nutzen könnten, wenngleich bei der Bahnfahrt derzeit ein Umsteigen erforderlich ist. Allerdings dauert die Fahrt mit dem Bus länger. Im Gegensatz zum Bus ist in der Bahn eine Toilette vorhanden, die jedoch verschmutzt ist. Beide öffentlichen Verkehrsmittel sind im Innenraum sauber. Im Bus steht allen Befragten für die ganze Fahrt ein bequemer Sitzplatz zur Verfügung, in der Bahn hingegen nur (zufällig ausgewählten) 50 % der Befragten für die ganze Fahrt, 25 % für die halbe Fahrt und 25 % gar nicht. Bei beiden Verkehrsmitteln wurde wie im ersten Szenario die

Bewertung des Konstrukts der *Informationen* der aktuellen bzw. alternativen Bahnfahrt als zusammenfassender Konstruktwert übernommen.⁴³⁰ Im sechsten Szenario weist die Gruppe der *nicht wahlfreien Bahnnutzer* Wahlanteile für die Bahn von 72,1 % und für den Bus von 27,9 % auf. Die Wahlanteile der Bahn, des Busses und des Pkw betragen bei der Gruppe der „*wahlfreien*“ Pkw-Nutzer 18,1 %, 17,5 % bzw. 64,4 %.

Das siebte Szenario unterscheidet sich vom sechsten Szenario bezüglich einer sauberen Toilette in der Bahn. Diese Veränderung führt bei den *nicht wahlfreien Bahnnutzern* zu einer Steigerung des Wahlanteils der Bahn um 2,02 und bei den „*wahlfreien*“ Pkw-Nutzern um eine Steigerung um 5,60 Prozentpunkte. Gleichzeitig sinken bei letzterer Gruppe die Wahlanteile des Busses und des Pkw um 2,17 bzw. 3,43 Prozentpunkte.

Eine Übersicht der Szenarien der Gruppe der *wahlfreien Bahnnutzer* ist in Tab. 7.32 und eine Übersicht der ermittelten Wahlanteile in Tab. 7.33 dargestellt.

Abweichend von den Szenarien der anderen beiden Nutzergruppen konnten in den ersten fünf Szenarien nicht durchgehend die zusammenfassenden Konstruktwerte für alle drei Konstrukte verwendet werden. Stattdessen wurden unterhalb des Asts, der sich auf ein Subexperiment bezieht (vgl. zum Aufbau des Modells Abb. 7.13), Ausprägungen für die jeweiligen Attribute dieses Subexperiments sowie zusammenfassende Konstruktwerte der anderen beiden Konstrukte verwendet. Darüber hinaus sind diese Szenarien mit den Szenarien der anderen beiden Nutzergruppen vergleichbar.

Die Auswahl der Ausprägungen der qualitativen Attribute erfolgte vor dem Hintergrund der durchschnittlich hohen Konstruktbewertung und der Analyse der Rating-Experimente, sodass überwiegend die oberste/beste Ausprägung der Attribute verwendet wurde. Mögliche Abweichungen zwischen diesen Attributsausprägungen und der zugehörigen Konstruktbewertung des Befragten wurden bei den ersten fünf Szenarien in Kauf genommen, da hier die Analyse der Änderungen von Wahlanteilen infolge

⁴³⁰ Die Verwendung des Konstrukts der *Informationen* anstelle der zugehörigen Attribute erfolgte aufgrund der Probleme bezüglich der Signifikanz der geschätzten Parameter.

Tab. 7.32: Szenarien der Simulation für die wahlfreien Bahnnutzer

| | Bahn | Bus |
|------------|---|--|
| Szenario 1 | <p>Ast der <i>Verbindungsqualität</i>:</p> <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> der aktuellen Fahrt mit der Bahn - <u>Pünktlichkeit</u>: <i>pünktlich bzw. max. 3 Minuten Verspätung</i> - <u>Umsteigehäufigkeit</u>: <i>1 x Umsteigen mit Anschlussgarantie</i> - <u>Takt</u>: <i>alle 30 Minuten</i> - zusammenfassende <u>Konstruktwerte</u> <u>Komfort</u> und <u>Informationen</u>: Konstruktbewertungen der aktuellen Fahrt (sofern angegeben, ansonsten Mediane der Konstruktbewertung) <p>Ast des <i>Komforts</i>:</p> <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> der aktuellen Fahrt mit der Bahn - <u>Toilette im Zug</u>: <i>sauber</i> - <u>Sauberkeit Zug innen</u>: <i>sauber</i> - <u>Sitzplatzverfügbarkeit im Zug</u>: <i>ganze Fahrt Sitzplatz</i> - (<u>Komfort der Sitzplätze</u>: Parameter nicht signifikant) - zusammenfassende <u>Konstruktwerte</u> <u>Verbindungsqualität</u> und <u>Informationen</u>: Konstruktbewertungen der aktuellen Fahrt (sofern angegeben, ansonsten Mediane der Konstruktbewertung) <p>Ast der <i>Informationen</i>:</p> <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> der aktuellen Fahrt mit der Bahn - (<u>Attribute des Konstrukts der Informationen</u>: Parameter nicht signifikant) - zusammenfassende <u>Konstruktwerte</u> <u>Verbindungsqualität</u> und <u>Komfort</u>: Konstruktbewertungen der aktuellen Fahrt (sofern angegeben, ansonsten Mediane der Konstruktbewertung) | <p>Ast der <i>Verbindungsqualität</i>:</p> <p>identische Ausprägungen wie Bahnalternative in Szenario 1</p> <p>Ast des <i>Komforts</i>:</p> <p>identische Ausprägungen wie Bahnalternative in Szenario 1</p> <p>Ast der <i>Informationen</i>:</p> <p>identische Ausprägungen wie Bahnalternative in Szenario 1</p> |
| Szenario 2 | <p>identische Ausprägungen wie Bahnalternative in Szenario 1, aber <u>Erhöhung des Fahrpreises</u> um 1 %</p> | <p>identische Ausprägungen wie Busalternative in Szenario 1</p> |

Fortsetzung nächste Seite

Fortsetzung Tab. 7.32

| | Bahn | Bus |
|------------|---|---|
| Szenario 3 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Erhöhung der Gesamtreisezeit</u> um 10 Minuten |
| Szenario 4 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Reduzierung des Fahrpreises</u> um 10 % |
| Szenario 5 | identische Ausprägungen wie Bahnalternative in Szenario 1 | identische Ausprägungen wie Busalternative in Szenario 1, aber <u>Erhöhung der Gesamtreisezeit</u> um 10 Minuten und <u>Reduzierung des Fahrpreises</u> um 10 % |
| Szenario 6 | <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> wie Bahnalternative in Szenario 1 - <u>Toilette im Zug</u>: verschmutzt - <u>Sauberkeit Zug innen</u>: sauber - <u>Sitzplatzverfügbarkeit im Zug</u>: ganze Fahrt Sitzplatz (50 %), halbe Fahrt Sitzplatz (25 %), kein freier Sitzplatz (25 %) - <u>(Komfort der Sitzplätze</u>: Parameter nicht signifikant) - zusammenfassende <u>Konstruktivwerte Verbindungsqualität</u> und <u>Informationen</u>: wie in Szenario 1 | <ul style="list-style-type: none"> - <u>Gesamtreisezeit</u> und <u>Fahrpreis</u> wie Busalternative in Szenario 3 (<u>Erhöhung der Gesamtreisezeit</u> um 10 Minuten) - <u>Toilette im Bus</u>: nicht vorhanden - <u>Sauberkeit Bus innen</u>: sauber - <u>Sitzplatzverfügbarkeit im Bus</u>: ganze Fahrt Sitzplatz - <u>(Komfort der Sitzplätze</u>: Parameter nicht signifikant) - zusammenfassende <u>Konstruktivwerte Verbindungsqualität</u> und <u>Informationen</u>: wie in Szenario 1 |
| Szenario 7 | identische Ausprägungen wie Bahnalternative in Szenario 6, aber <u>Toilette im Zug</u> : sauber | identische Ausprägungen wie Busalternative in Szenario 6 |

Quelle: Eigene Darstellung.

einer Änderung der Gesamtreisezeit und/oder des Fahrpreises jeweils eines öffentlichen Verkehrsmittels im Vordergrund steht.

Im ersten Szenario betragen bei der Gruppe der *wahlfreien Bahnnutzer* die Wahlanteile der Bahn, des Busses und des Pkw 60,9 %, 22,8 % bzw. 16,3 %. Im Vergleich dazu führt eine Erhöhung des Fahrpreises der Bahn

um 1 % (Szenario 2) zu einem Rückgang des Wahlanteils der Bahn um 0,38 Prozentpunkte, sodass die Preiselastizität $(-0,38 \%/60,93 \%) / 1 \% = -0,63$ (unelastische Nachfrage) beträgt. Der Wahlanteil des Busses steigt um 0,31 Prozentpunkte und der Wahlanteil des Pkw um 0,08 Prozentpunkte.

Eine Erhöhung der Gesamtreisezeit des Busses um 10 Minuten (Szenario 3) führt im Vergleich zum ersten Szenario zu einer Senkung des Wahlanteils des Busses um 11,15 Prozentpunkte und zu einer Erhöhung der Wahlanteile der Bahn und des Pkw um 10,43 bzw. 0,72 Prozentpunkte.

Durch eine Senkung des Fahrpreises des Busses um 10 % (Szenario 4) reduzieren sich die Wahlanteile der Bahn und des Pkw um 3,30 bzw. 0,32 Prozentpunkte, wohingegen sich der Wahlanteil des Busses um 3,62 Prozentpunkte erhöht.

Bei einer Erhöhung der Gesamtreisezeit des Busses um 10 Minuten und einer gleichzeitigen Preissenkung um 10 % (Szenario 5) sinkt der Wahlanteil des Busses um 8,94 Prozentpunkte, während die Wahlanteile der Bahn und des Pkw um 8,40 bzw. 0,54 Prozentpunkte steigen.

Abweichend von den ersten fünf Szenarien wurden im sechsten Szenario – um mögliche Abweichungen zwischen Attributsausprägungen und zusammenfassenden Konstrukten zu verhindern – bei der Bestimmung der individuenspezifischen Wahlwahrscheinlichkeiten ausschließlich die Parameter des Asts eines Subexperiments, in diesem Fall des *Komforts*, verwendet.⁴³¹ Der Bus- und der Bahnalternative wurden unterschiedliche hypothetische Ausprägungen der Attribute mit Bezug zum *Komfort* zugeordnet; diese entsprechen den Ausprägungen im sechsten Szenario der anderen beiden Nutzergruppen. Abweichend wurde jedoch nicht nur für das Konstrukt der *Informationen*, sondern auch für das Konstrukt der *Verbindungsqualität* die Bewertung der aktuellen bzw. alternativen Bahnfahrt übernommen.⁴³² Darüber hinaus unterscheidet sich dieses Szenario

⁴³¹ Die elementaren Alternativen unterhalb des Asts der *Verbindungsqualität* bzw. der *Informationen* wurden als durchgehend nicht verfügbar modelliert. In die Berechnung der individuenspezifischen Wahlwahrscheinlichkeiten wurden die Fahrten aller *wahl-freien Bahnnutzer* einbezogen, d. h. die elementaren Alternativen des Asts des *Komforts* wurden durchgehend verwendet.

⁴³² Die Verwendung der Konstrukte anstelle der zugehörigen Attribute erfolgte modellbeding.

Tab. 7.33: Mittelwert der Wahlwahrscheinlichkeiten der Alternativen der wahlfreien Bahnnutzer

| | Wahlfreie Bahnnutzer | | |
|------------|----------------------|--------|--------|
| | Bahn | Bus | Pkw |
| Szenario 1 | 60,9 % | 22,8 % | 16,3 % |
| Szenario 2 | 60,5 % | 23,1 % | 16,4 % |
| Szenario 3 | 71,4 % | 11,6 % | 17,0 % |
| Szenario 4 | 57,6 % | 26,4 % | 16,0 % |
| Szenario 5 | 69,3 % | 13,8 % | 16,8 % |
| Szenario 6 | 57,6 % | 21,9 % | 20,5 % |
| Szenario 7 | 66,8 % | 14,7 % | 18,5 % |

Quelle: Eigene Berechnungen.

nicht von dem sechsten Szenario der anderen beiden Nutzergruppen. Die Wahlanteile der Bahn, des Busses und des Pkw betragen in diesem sechsten Szenario 57,6 %, 21,9 % bzw. 20,5 %.

Auch bei dieser Nutzergruppe unterscheidet sich das siebte Szenario vom sechsten Szenario bezüglich einer sauberen Toilette in der Bahn. Dadurch steigt der Wahlanteil der Bahn um 9,21 Prozentpunkte und die Wahlanteile des Busses und des Pkw sinken um 7,16 bzw. 2,05 Prozentpunkte.

Auf gleiche Weise können ausschließlich die Parameter des Asts der *Verbindungsqualität* (bzw. prinzipiell auch die Parameter des Asts der *Informationen*) verwendet werden, um die Wahlanteile in Abhängigkeit der Attribute der *Verbindungsqualität* (bzw. der *Informationen*) zu bestimmen.

Derartige exemplarische Simulationen können als Ausgangspunkt für Handlungsempfehlungen dienen. Dabei ist allerdings zu beachten, dass für eine valide Simulation die Kenntnis der wahrgenommenen Qualität erforderlich ist. Weist beispielsweise ein Verkehrsmittel in der Realität bereits eine hohe Qualität auf und wird diese im Ausgangsszenario als zu niedrig angegeben, kann dies dazu führen, dass durch die de facto geringere Qualitätsverbesserung der vorhergesagte Wahlanteil dieser Alternative überschätzt wird.

In Verbindung mit dem Mengengerüst eines Gesamtverkehrsmodells kann bei Kenntnis der wahrgenommenen Qualität der öffentlichen Verkehrsmittel eine Veränderung von Fahrgastzahlen infolge einer Veränderung der Qualität simuliert werden. Dadurch ließen sich beispielsweise Aus-

sagen treffen, in Bezug auf welches Attribut die Qualität des regionalen ÖPNV gesteigert werden müsste, um eine lenkungspolitisch motivierte Verkehrsverlagerung zu erreichen. Im Sinne eines effizienten Einsatzes öffentlicher Mittel sollten dabei die mit den Maßnahmen zur Steigerung der Qualität verbundenen Kosten den zusätzlichen Fahrgelderlösen, die aus einer Erhöhung des Fahrpreises und/oder aus erhöhten Fahrgastzahlen resultieren, gegenübergestellt werden. Ebenso ließen sich insbesondere bei enger werdenden Finanzspielräumen durch die öffentliche Hand⁴³³ die Auswirkungen infolge einer Senkung der Qualität analysieren.

7.14 Zusammenfassung

Aufbauend auf den Ergebnissen der Vorstudie wurde ein Erhebungsinstrument entwickelt, das zur Befragung von Nutzern und Nicht-Nutzern des regionalen SPNV in Zügen des Regionalverkehrs bzw. in Behörden (insb. Kfz-Zulassungsstellen) eingesetzt wurde. Um einen Bezug zu einer konkreten Fahrt herstellen zu können, wurden insbesondere Fahrtkosten und Gesamtreisezeit für die tatsächlich durchgeführte Fahrt sowie ggf. für ebendiese Fahrt unter Verwendung eines anderen Verkehrsmittels (Pkw bzw. Regionalbahn) erhoben. In Rating-Experimenten wurden die Befragten mit den drei Konstrukten *Verbindungsqualität*, *Komfort* und *Informationen*, den zugehörigen Qualitätsattributen sowie der verwendeten Skala vertraut gemacht und es wurden darüber hinaus Konstruktbewertungen in Bezug auf die durchgeführte Fahrt erhoben. In den Subexperimenten der integrierten hierarchischen Informationsintegration konnten alle Befragten zwischen den beiden Alternativen Regionalbahn und Regionalbus wählen, wohingegen eine Pkw-Alternative nur bei realer Pkw-Verfügbarkeit wählbar war. Alle drei Alternativen waren durch die Ausprägungen der Attribute *Gesamtreisezeit* und *Fahrpreis/Fahrtkosten* charakterisiert, die auf Basis der Eigenschaften der beschriebenen Fahrt variiert wurden. Darüber hinaus wurden die beiden ÖV-Alternativen durch Ausprägungen der Attribute eines Konstrukts sowie hypothetische, zusammenfassende Konstruktwerte der beiden verbleibenden Konstrukte beschrieben.

⁴³³ Vgl. *ÖPNV-Zukunftskommission NRW* (2012), S. 5.

Zunächst wurden in die Analyse Datensätze von über 2500 Befragten einbezogen, die je nach Erhebungsort und Pkw-Verfügbarkeit als *nicht-wahlfreie Bahnnutzer* (Befragte in den Zügen ohne verfügbaren Pkw), *wahlfreie Bahnnutzer* (Befragte in den Zügen mit verfügbarem Pkw) und *Pkw-Nutzer* (Befragte in den Behörden mit verfügbarem Pkw) bezeichnet wurden. Die Rating-Experimente wurden anhand einer multiplen linearen Regressionsanalyse differenziert nach Nutzergruppen ausgewertet. Dabei wurde festgestellt, dass die Befragten bei der Konstruktbewertung auf Veränderungen der Attributsausprägungen reagiert haben, sodass ein grundsätzlicher Zusammenhang zwischen den Attributen und Konstrukten gegeben ist.

Für die drei Subexperimente wurden differenziert nach den Nutzergruppen der *nicht wahlfreien Bahnnutzer*, der *wahlfreien Bahnnutzer* und der *Pkw-Nutzer* multinomiale Logit-Modelle geschätzt. Für die Modelle der Subexperimente der *Verbindungsqualität* und des *Komforts* konnten bis auf wenige Ausnahmen durchgehend signifikante Parameter mit plausiblen Vorzeichen geschätzt werden, d. h. die Verkehrsmittelwahl wird durch die Ausprägungen der Attribute bzw. Konstrukte beeinflusst. Abweichend davon konnten in den Modellen des Subexperiments der *Informationen* für einige Qualitätsattribute keine signifikanten Parameter geschätzt werden. Für das Konstrukt der *Verbindungsqualität* wurde durchgehend der größte und für das Konstrukt der *Informationen* der kleinste Parameter geschätzt.

Bei allen drei Subexperimenten war bei der Gruppe der *Pkw-Nutzer* ein deutlicher Anteil von Befragten erkennbar, die bei den Wahlexperimenten stets die gleiche Alternative, in der Regel den Pkw, gewählt haben, d. h. diese Befragten haben nicht auf Veränderungen der Ausprägungen der Attribute bzw. Konstrukte reagiert. Durch einen Ausschluss der Daten dieser „*nicht wahlfreien*“ *Pkw-Nutzer* aus der weiteren Analyse und eine erneute Modellschätzung für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer*, d. h. für die Teilgruppe der *Pkw-Nutzer*, die zwischen den Alternativen gewechselt hat, konnte die Signifikanz der Nutzenparameter verbessert werden.

Zur Überprüfung der Gleichheit der Nutzenfunktionen der *nicht wahlfreien Bahnnutzer* und der *wahlfreien Bahnnutzer* sowie der *wahlfreien*

Bahnnutzer und der „*wahlfreien*“ *Pkw-Nutzer* wurden differenziert nach den drei Subexperimenten zweistufige Likelihood-Ratio-Tests durchgeführt. Die Hypothese der Gleichheit der Nutzenfunktionen wurde für alle Gruppen durchgehend bereits in der ersten Stufe des Testverfahrens abgelehnt. Daher erschien es erforderlich, die weitere Analyse getrennt nach Nutzergruppen durchzuführen.

Für die Gruppen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* (trimodale Modelle) wurde die dem multinomialen Logit-Modell zugrunde liegende Annahme der Unabhängigkeit von irrelevanten Alternativen anhand des Hausman-McFadden-Tests überprüft. Die Testergebnisse zeigten, dass die Annahme der Unabhängigkeit von irrelevanten Alternativen für die Gruppe der *wahlfreien Bahnnutzer* in allen drei Modellen und für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* im Modell der *Informationen* verletzt ist, sodass ein Nested Logit-Modell für die Analyse der Daten dieser beiden Nutzergruppen geeigneter erschien.

Das Nested Logit-Modell stellte bis auf eine Ausnahme durchgehend eine signifikante Verbesserung (in Bezug auf den Log-Likelihood-Wert) im Vergleich zum multinomialen Logit-Modell dar. Hinsichtlich der Signifikanz der Parameter wurden vergleichbare Ergebnisse wie bei den MNL-Modellen festgestellt.

Die Verwendung des NL-Modells anstelle des MNL-Modells erforderte eine erneute Überprüfung der Gleichheit der Nutzenfunktionen der *nicht wahlfreien* und der *wahlfreien Bahnnutzer* sowie der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer*, die ebenfalls differenziert nach den drei Subexperimenten anhand eines zweistufigen Likelihood-Ratio-Tests durchgeführt wurde. Zwischen den Nutzenfunktionen der *nicht wahlfreien* und der *wahlfreien Bahnnutzer* im Modell der *Verbindungsqualität* sowie zwischen denen der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzern* in allen drei Subexperimenten wurden Unterschiede festgestellt.

Nachdem mit dem multinomialen Logit-Modell für die Gruppe der *nicht wahlfreien Bahnnutzer* und mit dem Nested Logit-Modell für die Gruppe der *wahlfreien Bahnnutzer* bzw. der „*wahlfreien*“ *Pkw-Nutzer* geeignete Modelle gefunden waren und Unterschiede zwischen den Nutzergruppen

jeweils innerhalb eines Subexperiments festgestellt wurden, erfolgte differenziert nach den drei Nutzergruppen die Überprüfung der Prozessgleichheit zwischen den Subexperimenten. Hierzu wurde ein zweistufiger Likelihood-Ratio-Test verwendet. Damit wurde zum einen getestet, ob die Modelle der unterschiedlichen Subexperimente zu jeweils einem gemeinsamen Modell vereinigt werden können, d. h. ob auf Basis der Daten der drei Subexperimente ein geeignetes gemeinsames Modell geschätzt werden kann. Zum anderen kann dies als eine Überprüfung des hierarchischen Wahlmodells für diese konkrete Anwendung angesehen werden.

Bei der Gruppe der *nicht wahlfreien Bahnnutzer* sowie der Gruppe der „*wahlfreien*“ *Pkw-Nutzer* bestehen keine Unterschiede zwischen den Nutzenfunktionen der drei Subexperimente. Es liegt Prozessgleichheit vor, d. h. unabhängig von dem eingesetzten Subexperiment werden die gleichen Wahlentscheidungen abgebildet. Hinsichtlich der Überprüfung des hierarchischen Wahlmodells an sich liegen somit zwei erfolgreiche Anwendungen vor. Die Modelle der Subexperimente können für diese beiden Gruppen jeweils zu einem gemeinsamen Modell vereint werden. Bei der Gruppe der *wahlfreien Bahnnutzer* besteht Gleichheit der Nutzenparameter bei unterschiedlichen Skalierungsparametern. Hinsichtlich der Überprüfung des hierarchischen Wahlmodells handelt es sich daher nur bedingt um eine erfolgreiche Anwendung. Die Modelle der Subexperimente können zu einem gemeinsamen Modell vereint werden, sofern die Unterschiede zwischen den Skalierungsparametern berücksichtigt werden.

In den gemeinsamen Modellen der Gruppen der *nicht wahlfreien Bahnnutzer*, der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* weisen die Attribute der *Verbindungsqualität* und des *Komforts* überwiegend signifikante Parameter mit plausiblen Vorzeichen auf. Die Probleme in Bezug auf die Signifikanz der Attribute der *Informationen* bestanden fort. Somit unterscheiden sich die gemeinsamen Modelle bezüglich der Signifikanz der Parameter nicht oder nur sehr geringfügig von den entsprechenden einzelnen Modellen. Auch die Rangfolge der Parameter der Konstrukte hat sich gegenüber den einzelnen Modellen nicht verändert, d. h. für das Konstrukt der *Verbindungsqualität* wurde der größte und für das Konstrukt der *Informationen* der kleinste Parameter geschätzt.

Aufbauend auf den Parametern der gemeinsamen Modelle wurden differenziert nach Nutzergruppen Zahlungsbereitschaften für die Veränderung der Ausprägungen der Qualitätsattribute bestimmt. Ein Vergleich der Zahlungsbereitschaften der drei Nutzergruppen zeigte fast durchgehend für die Gruppe der „*wahlfreien*“ *Pkw-Nutzer* die größten und für die Gruppe der *nicht wahlfreien Bahnnutzer* die niedrigsten Zahlungsbereitschaften. Für Veränderungen des Konstrukts der *Verbindungsqualität* liegt die höchste und für Veränderungen des Konstrukts der *Informationen* die niedrigste Zahlungsbereitschaft vor.

Die vergleichsweise hohe Zahlungsbereitschaft für die *Verbindungsqualität* wird auch bei den Zahlungsbereitschaften für die Ausprägungen der einzelnen Attribute dieses Konstrukts deutlich. Die Attribute *Pünktlichkeit*, *Umsteigen* und *Takt* gehören zu den vier qualitativen Attributen mit der größten Spannweite zwischen der Zahlungsbereitschaft für die oberste/beste Ausprägung und der Zahlungsbereitschaft für die unterste/schlechteste Ausprägung. Derartige Vergleiche können als Indikator für die Wichtigkeit der Attribute dienen, sind jedoch stets abhängig vom Aufbau des Wahlexperiments.

Des Weiteren wurde grundsätzlich gezeigt, wie sich die Steigerung eines Qualitätsattributs mit der Senkung anderer Qualitätsattribute kompensieren lässt. Eine nutzenneutrale Veränderung für eine Gruppe kann jedoch für eine andere Gruppe nicht nutzenneutral sein, sodass eine Änderung der Wahlwahrscheinlichkeiten erfolgt.

Differenziert nach Nutzergruppen wurde deshalb schließlich eine Simulation von Wahlanteilen in Abhängigkeit der Ausprägungen der Attribute bzw. der Konstrukte für unterschiedliche Szenarien exemplarisch durchgeführt. Dabei wurden die Gesamtreisezeiten und Fahrpreise bzw. Fahrtkosten von den real durchgeführten Fahrten sowie die zugehörigen Konstruktbewertungen als Simulationsgrundlage verwendet. Für die Gruppen der *nicht wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* können die Ausprägungen der Attribute unterschiedlicher Konstrukte gleichzeitig variiert werden. Aufgrund der Unterschiede zwischen den Skalierungsparametern im gemeinsamen Modell der *wahlfreien Bahnnutzer* können hingegen grundsätzlich nur die Ausprägungen der Attribute eines Konstrukts verändert werden.

Aufbauend auf einer derartigen Simulation können bei Kenntnis der wahrgenommenen Qualität der öffentlichen Verkehrsmittel in Verbindung mit dem Mengengerüst eines Gesamtverkehrsmodells Auswirkungen auf die Fahrgastzahlen infolge einer Veränderung der Qualität simuliert und unter Berücksichtigung von Kostenaspekten ökonomische Handlungsempfehlungen für die Praxis abgeleitet werden.

8 Fazit und Ausblick

Die der vorliegenden Arbeit zugrunde liegenden Ziele, die zum Ersten darin bestanden, den Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr zu analysieren, und zum Zweiten darin, die hierarchische Struktur des dabei eingesetzten Verfahrens der integrierten hierarchischen Informationsintegration für diese konkrete Anwendung zu überprüfen, wurden erreicht.

Zur Analyse des Zusammenhangs zwischen Qualität und Verkehrsmittelwahl wurden Discrete Choice Experimente durchgeführt, die auf dem Verfahren der integrierten hierarchischen Informationsintegration beruhen. Dabei wurden insgesamt 13 Qualitätsattribute – entweder direkt als Attribute oder indirekt durch die zusammenfassenden Konstruktwerte – berücksichtigt. In allen Verkehrsmittelwahlmodellen erwiesen sich die Parameter der Attribute *Gesamtreisezeit* und *Fahrpreis/Fahrtkosten* durchgehend und die Parameter der Attribute der *Verbindungsqualität* und des *Komforts* überwiegend als signifikant. Daraus kann gefolgert werden, dass diese Attribute einen Einfluss auf die Verkehrsmittelwahl ausüben. Die Parameter der Attribute der *Informationen* erwiesen sich überwiegend als nicht signifikant. Allerdings kann daraus nicht gefolgert werden, dass diese Attribute keinen Einfluss auf die Verkehrsmittelwahl ausüben.

Die Ursachen für die nicht signifikanten Parameter der Attribute der *Informationen* können vielseitig sein. Da für das Konstrukt der *Informationen* in den Modellen der anderen Subexperimente durchgehend der kleinste Parameter geschätzt wurde, waren auch für die Attribute der *Informationen* vergleichsweise kleine Parameter zu erwarten. Allgemein erfordern Attribute, die nur einen geringen Einfluss auf die Verkehrsmittelwahl ausüben, eine größere Datenbasis.

Bei der Datenerhebung im Subexperiment der *Informationen* erfolgte eine Unterscheidung nach den Arten der Informationsvermittlung (Anzeige bzw. Durchsage), deren Kombination die vier Attributsausprägungen bildeten. Dadurch erhöhte sich die Anzahl der zu schätzenden Parameter und auch die Komplexität der Wahlentscheidung für den Befragten. Mög-

licherweise wären bessere Schätzergebnisse erzielt worden, wenn hier nur dahingehend unterschieden worden wäre, ob überhaupt die Informationen bereitgestellt werden. Ein weiterer Aspekt könnte sein, dass einige Befragte eher Anzeigen bevorzugen, andere hingegen Durchsagen. Auch könnten Informationen allgemein für Befragte, die den ÖPNV eher selten nutzen, von größerer Bedeutung sein als für ÖV-Berufspendler. Neben einem derartigen nicht homogenen Entscheidungsverhalten könnte auch eine teilweise Nicht-Beachtung der Attribute (*attribute non-attendance*) ursächlich für die nicht signifikanten Parameter sein.

Sowohl das multinomiale Logit-Modell als auch das Nested Logit-Modell beruhen auf der Annahme der Verhaltenshomogenität. Bei der Analyse wurde bereits zwischen den Nutzergruppen der *nicht wahlfreien Bahnnutzer*, der *wahlfreien Bahnnutzer* und der „*wahlfreien*“ *Pkw-Nutzer* unterschieden. Darüber hinaus könnte eine weitere Segmentierung zu homogeneren Gruppen führen, wodurch die Varianz der Nutzenparameter reduziert und die Signifikanz der Parameter der *Informationen* verbessert werden könnten. Dem steht jedoch eine geringere Datenbasis innerhalb der Gruppe entgegen.

Auch Modelle, bei denen die Restriktion der Verhaltenshomogenität nicht besteht, wie beispielsweise das Mixed Logit Modell oder das Latent Class Modell, könnten zur Analyse dieser Wahlentscheidungen eingesetzt werden. Dabei wäre für jeden einzelnen Parameter die zugrunde liegende Verteilungsfunktion zu bestimmen. Zur Überprüfung der Prozessgleichheit zwischen den Subexperimenten kann allerdings nicht der zweistufige Likelihood-Ratio-Test eingesetzt werden, sodass diesbezüglich andere Testverfahren⁴³⁴ erforderlich sind.

Das Verfahren der HII-I wurde im Verkehrsbereich bislang noch nicht eingesetzt, sondern nur das Verfahren der HII, bei dem die zugrunde liegende hierarchische Struktur nicht überprüft werden kann. Aber auch außerhalb des Verkehrsbereichs gibt es nur wenige Anwendungen der HII-I, in denen die hierarchische Struktur getestet wurde. *Oppewal/Louviere/Timmermans* (1994) konnten für eine Anwendung der HII-I die Gleichheit der Nutzenparameter zwischen den Subexperimenten nicht

⁴³⁴ Vgl. für ein mögliches Testverfahren *Hensher/Rose/Greene* (2008) und *Hensher* (2008).

bestätigen. Für eine andere Anwendung stellten *van de Vyvere/Oppewal/Timmermans* (1998) zwar die Gleichheit der Nutzenparameter fest, doch bestanden Unterschiede zwischen den Skalierungsparametern, sodass diese Anwendung der HII-I nur als bedingt erfolgreich angesehen werden kann. Bei diesen Anwendungen basierte die Entwicklung der Konstrukte auf einer Literaturrecherche und Experteninterviews. Eine empirische Entwicklung der Konstrukte führten *Bos* (2004), *Bos et al.* (2004) und *Bos/Molin* (2006) durch, doch aufgrund der Verwendung der HII war dabei keine Überprüfung der Prozessgleichheit möglich.

In der vorliegenden Arbeit war die Anwendung der HII-I in Bezug auf zwei Nutzergruppen – aufgrund der Gleichheit der Nutzenparameter sowie der Skalierungsparameter – erfolgreich und in Bezug auf die dritte Nutzergruppe – aufgrund der Gleichheit der Nutzenparameter und der Ungleichheit der Skalierungsparameter – bedingt erfolgreich. Dieses könnte dadurch begünstigt worden sein, dass die empirische Entwicklung der Konstrukte möglicherweise dazu beigetragen hat, die Zuordnung der Attribute zu den Konstrukten in den Experimenten der HII-I für die Befragten verständlicher zu gestalten.

Durch diese Arbeit konnte ein Beitrag für die empirische Evidenz des Verfahrens der HII-I geleistet werden. Es besteht jedoch noch weiterer Forschungsbedarf. Sofern die Prozessgleichheit durch die Zuordnung der Attribute zu den Konstrukten beeinflusst wird, ließe sich die Notwendigkeit einer empirischen Entwicklung der Konstrukte begründen. Des Weiteren könnte die Kenntnis, inwiefern unterschiedlich komplexe Subexperimente die Unterschiede in den Skalierungsparametern verursachen, ggf. bei der Gestaltung der Experimente der HII-I hilfreich sein, um eine hinsichtlich der Prozessgleichheit erfolgreiche Anwendung durchzuführen. Schließlich können dann im Rahmen einer Simulation der Wahlwahrscheinlichkeiten die Ausprägungen aller Attribute variiert werden.

Wenngleich noch weiterer Forschungsbedarf besteht, ist der im Rahmen dieser Arbeit analysierte Zusammenhang zwischen der Qualität und der Verkehrsmittelwahl je nach verwendeter Vertragsform für den Aufgabenträger und/oder für das Verkehrsunternehmen und damit aus gemeinwirtschaftlicher und/oder einzelwirtschaftlicher Sicht von Bedeutung. Aufbauend auf den Ergebnissen dieser Arbeit können in Verbindung mit dem

Mengengerüst eines Gesamtverkehrsmodells und unter Berücksichtigung von Kostenaspekten die Auswirkungen auf die Fahrgastzahlen im regionalen ÖPNV infolge einer Veränderung der Qualität simuliert werden. Dadurch können beispielsweise Handlungsempfehlungen zur Erreichung lenkungs- politisch motivierter Verkehrsverlagerungen abgeleitet werden.

Anhang

A. Skalenniveau der Daten der aggregierten Unähnlichkeitsmatrix

Gemäß *Stevens* (1946) wird hinsichtlich des Skalenniveaus zwischen Nominal-, Ordinal-, Intervall- und Ratioskala unterschieden.⁴³⁵ Bei der Intervall- und der Ratioskala handelt es sich um metrische Skalen.

In Bezug auf das Skalenniveau der Daten der Unähnlichkeitsmatrix gilt es zunächst zu prüfen, ob ein metrisches Skalenniveau vorliegt. Dieses ist gegeben, wenn die Axiome des metrischen Raumes

$$\rho(q, r) \geq 0 \text{ und } \rho(q, r) = 0 \text{ genau dann, wenn } q = r \quad (\text{A.1})$$

(Nichtnegativität)

$$\rho(q, r) = \rho(r, q) \quad (\text{A.2})$$

(Symmetrie)

$$\rho(q, r) \leq \rho(q, s) + \rho(s, r) \quad (\text{A.3})$$

(Dreiecksungleichung)

erfüllt sind.⁴³⁶

Zunächst lässt sich für die Unähnlichkeitsmatrix auf Individualebene, welche mit \mathbf{M} bezeichnet wird, veranschaulichen, dass diese drei Axiome erfüllt sind.⁴³⁷ Die Zelle c_{qr} der Unähnlichkeitsmatrix enthält den Wert 0, wenn die beiden Objekte q und r in die gleiche Gruppe sortiert wurden und den Wert 1, wenn sie in unterschiedliche Gruppen sortiert wurden. Das erste Axiom ist somit erfüllt, denn zum einen liegt $\rho(q, r) = 0$ nur bei Gleichheit vor, zum anderen enthält die Zelle c_{qr} nie einen negativen Wert.

⁴³⁵ Die Klassifikation nach *Stevens* (1946) gilt als die am weitesten verbreitete Klassifikation des Skalenniveaus, vgl. *Schnell/Hill/Esser* (2011), S. 134–135.

⁴³⁶ Vgl. *Bronstein* (1999), S. 602.

⁴³⁷ Vgl. *Miller* (1969), S. 171–174 und *Green/Carmone/Smith* (1989), S. 28–29.

Alle Werte beziehen sich auf Objektpaare. Da sich das Objekt q zu r wie das Objekt r zu q verhält, sind $\rho(q,r)$ und $\rho(r,q)$ identisch. Daraus folgt, dass die Unähnlichkeitsmatrix symmetrisch ist.

Bezüglich des dritten Axioms lassen sich zwei Ausgangssituationen unterscheiden:

Bei der ersten Ausgangssituation befinden sich q und r in einer Gruppe, somit gilt $\rho(q,r) = 0$. Das Objekt s befindet sich entweder ebenfalls in dieser Gruppe, dann gilt $\rho(q,s) = 0$ und $\rho(s,r) = 0$ oder in einer anderen Gruppe, dann gilt $\rho(q,s) = 1$ und $\rho(s,r) = 1$. Mit $0 \leq 0$ bzw. $0 \leq 2$ ist die Dreiecksungleichung bei dieser Ausgangssituation erfüllt.

Bei der zweiten Ausgangssituation befinden sich q und r in unterschiedlichen Gruppen, somit gilt $\rho(q,r) = 1$. Das Objekt s befindet sich entweder in der gleichen Gruppe wie q , dann gilt $\rho(q,s) = 0$ und $\rho(s,r) = 1$, in der gleichen Gruppe wie r , dann gilt $\rho(q,s) = 1$ und $\rho(s,r) = 0$ oder in einer weiteren Gruppe, dann gilt $\rho(q,s) = 1$ und $\rho(s,r) = 1$. Mit $1 \leq 1$ bzw. $1 \leq 2$ ist die Dreiecksungleichung auch bei dieser Ausgangssituation gültig. Somit ist das dritte Axiom erfüllt.

Da alle drei Axiome erfüllt sind, genügen die Daten einer Metrik. Da die Summe einer Metrik ebenfalls eine Metrik ist, bleibt das metrische Skalenniveau erhalten, wenn die Unähnlichkeitsmatrizen auf Individual-ebene durch Summation aggregiert werden.

Demzufolge weist die aggregierte Unähnlichkeitsmatrix metrisches, mindestens intervallskaliertes Skalenniveau auf. Die Ratioskala unterscheidet sich von der Intervallskala durch die Möglichkeit, einen „natürlichen Nullpunkt“ abzubilden, der die Abwesenheit eines gemessenen Merkmals anzeigt.⁴³⁸ Ein solcher „natürlicher Nullpunkt“ ist bei der aggregierten Unähnlichkeitsmatrix gegeben und liegt vor, wenn alle Befragten ein Objektpaar in die gleiche Gruppe sortiert haben. Somit weisen die Daten der aggregierten Unähnlichkeitsmatrix ein ratio-skaliertes Skalenniveau auf.

⁴³⁸ Vgl. Schnell/Hill/Esser (2011), S. 135–136.

C. Erhebungsinstrument zur empirischen Entwicklung der Konstrukte

Abb. C.1: Screenshot des ersten Fragebogenteils mit Fragen zum allgemeinen Verkehrsverhalten

The screenshot shows a web browser window titled 'F1Einleitung'. The page header includes the logo of Fachhochschule Gelsenkirchen, the title 'QUASIMIR Qualitätssimulationsmodell im Regionalverkehr', and the logo of the Bundesministerium für Bildung und Forschung. The main content area contains the following elements:

- Interviewernummer:
- Erhebungsort:
- Greeting: 'Schönen guten Tag!'
- Introduction: 'Wir führen eine Studie zur Wahrnehmung von Qualitätseigenschaften im öffentlichen Personennahverkehr (ÖPNV) durch. Ihre Teilnahme an der Befragung dauert ca. 15 Minuten.'
- Question: 'Zu welchem Zweck führen Sie hauptsächlich Fahrten durch?' with a dropdown menu.
- Question: 'Besitzen Sie einen Pkw-Führerschein?' with radio buttons for 'ja' and 'nein'.
- Question: 'Steht Ihnen ein Pkw zur Verfügung?' with radio buttons for 'immer', 'meistens', 'manchmal', 'selten', and 'nie'.
- Question: 'Besitzen Sie eine Monatskarte?' with radio buttons for 'ja' and 'nein'.
- Question: 'Wie oft fahren Sie mit dem Zug?' with radio buttons for 'mehrmals in der Woche', 'mehrmals im Monat', 'selten', and 'nie'.
- Question: 'Wie oft nutzen Sie den Pkw?' with radio buttons for 'mehrmals in der Woche', 'mehrmals im Monat', 'selten', and 'nie'.
- 'weiter' button

The bottom of the browser window shows a status bar with 'Datensatz: 1 von 1', 'Kein Filter', and 'Suchen'.

Quelle: Eigene Darstellung.

Abb. C.2: Screenshot des zweiten Fragebogenteils mit Sortierungsaufgabe

The screenshot shows a Mozilla Firefox browser window displaying a web page from Fachhochschule Gelsenkirchen. The page title is 'Anleitung | Kommentar schreiben'. The main content area is a sorting task with the following elements:

- A list of 29 elements to be sorted, including: Umsteigehäufigkeit, Unterstellplatz / Wartehaus am Bahnsteig, Fahrkartenautomat im Zug, Sauberkeit der Toilette im Zug, Frequenz, Sauberkeit Zug außen, Anschlussgarantie, kein Fahrtenausfall, Videoüberwachung im Zug, Fahrkartenautomat am Bahnsteig, Fahrplanauskunft im Internet, Komfort der Sitzplätze, Informationsanzeige im Zug zu Anschlusszügen, Durchsage im Zug zu Anschlusszügen, and Abstellmöglichkeiten im Zug.
- Three sorting boxes with labels: 'Sauberkeit Zug innen', 'Gesamtreisezeit', and 'Fahrpreis'. Each box has a 'Zum Benennen anklicken' button above it.
- A 'Speichern' button at the bottom right.

Quelle: Eigene Darstellung.

Abb. C.3: Screenshot des dritten Fragebogenteils mit Bewertungsaufgaben (1 von 4)

F2aBewertungen

Fachhochschule Gelsenkirchen

QUASIMIR
Qualitätssimulationsmodell im
Regionalverkehr

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Wie wichtig sind für Sie folgende Qualitätseigenschaften in Bezug auf Ihre Verkehrsmittelwahl? 3/7

| | sehr unwichtig | | | | | sehr wichtig | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Kein Fahrtenausfall | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Pünktlichkeit | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Gesamtreisezeit (von Tür zu Tür) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Frequenz (Zahl der angebotenen Fahrten pro Stunde) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Umsteigehäufigkeit | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Anschlussgarantie | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrpreis | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sitzplatzverfügbarkeit im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Komfort der Sitzplätze | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Abstellmöglichkeiten im Zug (z.B. für Gepäck, Fahrräder, Kinderwagen u.ä.) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

weiter

Datensatz: 1 von 1 | Kein Filter | Suchen

Quelle: Eigene Darstellung.

Abb. C.4: Screenshot des dritten Fragebogenteils mit Bewertungsaufgaben (2 von 4)

F2bBewertungen

Fachhochschule Gelsenkirchen

QUASIMIR
Qualitätssimulationsmodell im
Regionalverkehr

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Wie wichtig sind für Sie folgende Qualitätseigenschaften in Bezug auf Ihre Verkehrsmittelwahl? 4/7

| | sehr unwichtig | | | | | sehr wichtig | |
|---|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Unterstellplatz / Wartehaus am Bahnsteig | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sauberkeit der Toilette im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sauberkeit Zug innen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sauberkeit Zug außen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrplaninformationen am Bahnsteig | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrplanauskunft im Internet | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Durchsagen im Zug bei Störungen oder Verspätungen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Durchsagen am Bahnsteig bei Störungen oder Verspätungen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Informationsanzeige am Bahnsteig bei Störungen oder Verspätungen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

weiter

Datensatz: 1 von 1 | Kein Filter | Suchen

Quelle: Eigene Darstellung.

Abb. C.5: Screenshot des dritten Fragebogenteils mit Bewertungsaufgaben (3 von 4)

F2cBewertungen

Fachhochschule Gelsenkirchen

QUASIMIR
Qualitätssimulationsmodell im Regionalverkehr

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Wie wichtig sind für Sie folgende Qualitätseigenschaften in Bezug auf Ihre Verkehrsmittelwahl? 5/7

| | sehr unwichtig | | | | | sehr wichtig | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Informationsanzeige am Zug (außen) | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Informationsanzeige im Zug zu Anschlusszügen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Durchsage im Zug zu Anschlusszügen | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Videoüberwachung im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Videoüberwachung am Bahnsteig | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sicherheitspersonal im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sicherheitspersonal am Bahnsteig | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Freundlichkeit des Personals | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrkartenautomat am Bahnsteig | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrkartenautomat im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

[weiter](#)

Datensatz: 1 von 1 | Kein Filter | Suchen

Quelle: Eigene Darstellung.

Abb. C.6: Screenshot des dritten Fragebogenteils mit Bewertungsaufgaben (4 von 4)

F2dBewertungen

Fachhochschule Gelsenkirchen

QUASIMIR
Qualitätssimulationsmodell im Regionalverkehr

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Wie wichtig sind für Sie folgende Qualitätseigenschaften in Bezug auf Ihre Verkehrsmittelwahl? 6/7

| | sehr unwichtig | | | | | sehr wichtig | |
|--|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Fahrkartenverkauf durch Personal am Bahnhof | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Fahrkartenverkauf durch Personal im Zug | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Verständlichkeit des Tarfsystems | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Sonstige Qualitätseigenschaften, die Ihre Verkehrsmittelwahl beeinflussen: | | | | | | | |
| <input type="text"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <input type="text"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| <input type="text"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

[weiter](#)

Datensatz: 1 von 1 | Kein Filter | Suchen

Quelle: Eigene Darstellung.

Abb. C.7: Screenshot des vierten Fragebogenteils mit soziodemografischen Fragen und Danksagung

The screenshot shows a web browser window titled 'F3Demografie'. The header includes the logo of Fachhochschule Gelsenkirchen, the project name 'QUASIMIR Qualitätssimulationsmodell im Regionalverkehr', and the logo of the Bundesministerium für Bildung und Forschung. The page number '7/7' is displayed in the top right corner. The main content area contains the following questions and input fields:

- Wie alt sind Sie?
- Ihr Geschlecht: männlich weiblich
- Wie viele Personen leben in Ihrem Haushalt?
- Wie viele dieser Personen sind unter 18 Jahren?

Below the questions is a thank-you message: 'Vielen Dank für Ihre Teilnahme an der Befragung!'. At the bottom right, there is a button labeled 'abschließen'. The footer of the browser window shows 'Datensatz: 1 von 1', 'Kein Filter', and a search button.

Quelle: Eigene Darstellung.

Abb. C.8: Screenshot des fünften Fragebogenteils für Anmerkungen des Interviewers

The screenshot shows a web browser window titled 'F4Anmerkungen'. The header is identical to the previous page, featuring the logos of Fachhochschule Gelsenkirchen, QUASIMIR, and the Bundesministerium für Bildung und Forschung. The page number '7/7' is not visible. The main content area contains the following text and input field:

- Anmerkungen (Interviewer):
-
- max. 255 Zeichen

At the bottom right, there is a button labeled 'abschließen'. The footer of the browser window shows 'Datensatz: 1 von 7', 'Kein Filter', and a search button.

Quelle: Eigene Darstellung.

D. Erhebungsinstrument mit HII-I für Bahnnutzer

Abb. D.1: Screenshot des ersten Fragebogenteils mit Anmeldung des Interviewers

Start_Export

Fachhochschule Gelsenkirchen

QUASIMIR - Qualitätssimulationsmodell im Regionalverkehr
Hauptuntersuchung

Version 1.8

GEFÖRDERT VOM
Bundesministerium für Bildung und Forschung

Interviewernummer:

Erhebungsort

Zur Datenerhebung

Exportieren

(speichert die Dateien im Ordner
C:\Users\jssv\Documents\QuasimirExporte\ab)

Quelle: Eigene Darstellung.

Abb. D.2: Screenshot des zweiten Fragebogenteils zur Beschreibung der aktuellen Fahrt

F1Einleitung

Schönen guten Tag!

Wir führen eine Studie zum Verkehrsmittelwahlverhalten durch.
Ihre Teilnahme an der Befragung dauert ca. 15 Minuten.

An welchem Bahnhof haben Sie diese Fahrt begonnen?

Wie haben Sie diesen Bahnhof erreicht?

An welchem Bahnhof endet für Sie diese Fahrt?

Wie erreichen Sie von dort aus Ihr Ziel?

Wie lange dauert diese Fahrt insgesamt? Minuten

Wie oft müssen Sie bei dieser Fahrt zwischen öffentlichen Verkehrsmitteln umsteigen?

Zu welchem Zweck führen Sie diese Fahrt durch? von nach

Wie hoch sind Ihre Fahrtkosten für diese Fahrt? €

Könnten Sie diese Fahrt auch mit dem Pkw durchführen? ja nein

Wie lange würde diese Fahrt mit dem Pkw insgesamt dauern? Minuten

Was würde Sie diese Fahrt insgesamt mit dem Pkw kosten? €

Weiter

Quelle: Eigene Darstellung.

Abb. D.3: Screenshot des dritten Fragebogens mit Rating-Experiment der Verbindungsqualität

Bitte bewerten Sie die Verbindungsqualität!

Pünktlichkeit pünktlich bzw. max. 3 Minuten Verspätung **3 bis 10 Minuten Verspätung** 10 bis 20 Minuten Verspätung

Umsteigehäufigkeit 0 x Umsteigen 1 x Umsteigen mit Anschlussgarantie **1 x Umsteigen ohne Anschlussgarantie**

Takt alle 30 Minuten **alle 60 Minuten** alle 120 Minuten

Bewertung +++++ ++++ +++ ++ + - -- --- ---- -----

Weiter

Quelle: Eigene Darstellung.

Abb. D.4: Screenshot des dritten Fragebogens mit Rating-Experiment des Komforts

Bitte bewerten Sie den Komfort!

Toilette im Zug sauber **verschmutzt** nicht vorhanden

Sauberkeit Zug innen **sauber** Boden klebrig, Papier liegt auf den Sitzen

Sitzplatzverfügbarkeit im Zug ganze Fahrt Sitzplatz halbe Fahrt Sitzplatz **kein freier Sitzplatz**

Komfort der Sitzplätze bequem **unbequem**

Bewertung +++++ ++++ +++ ++ + - -- --- ---- -----

Weiter

Quelle: Eigene Darstellung.

Abb. D.5: Screenshot des dritten Fragebogenteils mit Rating-Experiment der Informationen

Bitte bewerten Sie die Informationen!

| | | |
|---|--|--|
| Fahrplaninformationen am Bahnsteig | <input type="radio"/> vorhanden | <input checked="" type="radio"/> nicht vorhanden |
| Informationen im Zug zu Grund und Dauer von Störungen | <input checked="" type="radio"/> Durchsage und Anzeige | <input type="radio"/> Durchsage |
| | <input type="radio"/> Anzeige | <input type="radio"/> keine Informationen |
| Informationen am Bahnhof zu Grund und Dauer von Störungen | <input type="radio"/> Durchsage und Anzeige | <input type="radio"/> Durchsage |
| | <input type="radio"/> Anzeige | <input checked="" type="radio"/> keine Informationen |
| Informationen im Zug zu aktuellen Anschlusszügen | <input type="radio"/> Durchsage und Anzeige | <input type="radio"/> Durchsage |
| | <input checked="" type="radio"/> Anzeige | <input type="radio"/> keine Informationen |

Bewertung: ++++++ +++++ ++++ +++ ++ + - -- --- ---- -----

Quelle: Eigene Darstellung.

Abb. D.6: Screenshot des dritten Fragebogenteils mit Bewertung der aktuellen Fahrt in Bezug auf das Konstrukt der Verbindungsqualität

Bitte bewerten Sie für DIESE Fahrt die Verbindungsqualität!

+++++ ++++ +++ ++ + - -- --- ---- -----

Quelle: Eigene Darstellung.

Abb. D.7: Screenshot des vierten Fragebogenteils mit trimodalem Subexperiment (HII-I) der Verbindungsqualität

WahlVerbindung

Wenn Sie bei DIESER Fahrt nur genau diese Verkehrsmittel nutzen könnten, womit würden Sie fahren?

| Bahn | Regionalbus | Pkw |
|---|---|---|
| 3 bis 10 Minuten Verspätung | 10 bis 20 Minuten Verspätung | Fahrtkosten 12,60 € |
| 1 x Umsteigen ohne Anschlussgarantie | 1 x Umsteigen mit Anschlussgarantie | Gesamtreisezeit 30 min |
| fährt alle 60 Minuten | fährt alle 120 Minuten | |
| Komfort + + + + | Komfort + + + + | |
| Informationen + + + + | Informationen - - | |
| Fahrpreis 8,80 € | Fahrpreis 8,40 € | |
| Gesamtreisezeit 35 min | Gesamtreisezeit 45 min | |
| <input type="button" value="Diese Alternative wählen"/> | <input type="button" value="Diese Alternative wählen"/> | <input type="button" value="Diese Alternative wählen"/> |

Quelle: Eigene Darstellung.

Abb. D.8: Screenshot des vierten Fragebogenteils mit trimodalem Subexperiment (HII-I) des Komforts

WahlKomfort

Wenn Sie bei DIESER Fahrt nur genau diese Verkehrsmittel nutzen könnten, womit würden Sie fahren?

| Bahn | Regionalbus | Pkw |
|---|---|---|
| Toilette ist verschmutzt | keine Toilette vorhanden | Fahrtkosten 10,80 € |
| Zug ist innen sauber | Bus ist innen sauber | Gesamtreisezeit 30 min |
| halbe Fahrt Sitzplatz | ganze Fahrt Sitzplatz | |
| Sitzplatz ist bequem | Sitzplatz ist bequem | |
| Verbindungsqualität + + | Verbindungsqualität + + + + | |
| Informationen + + | Informationen + + | |
| Fahrpreis 8,80 € | Fahrpreis 8,40 € | |
| Gesamtreisezeit 40 min | Gesamtreisezeit 35 min | |
| <input type="button" value="Diese Alternative wählen"/> | <input type="button" value="Diese Alternative wählen"/> | <input type="button" value="Diese Alternative wählen"/> |

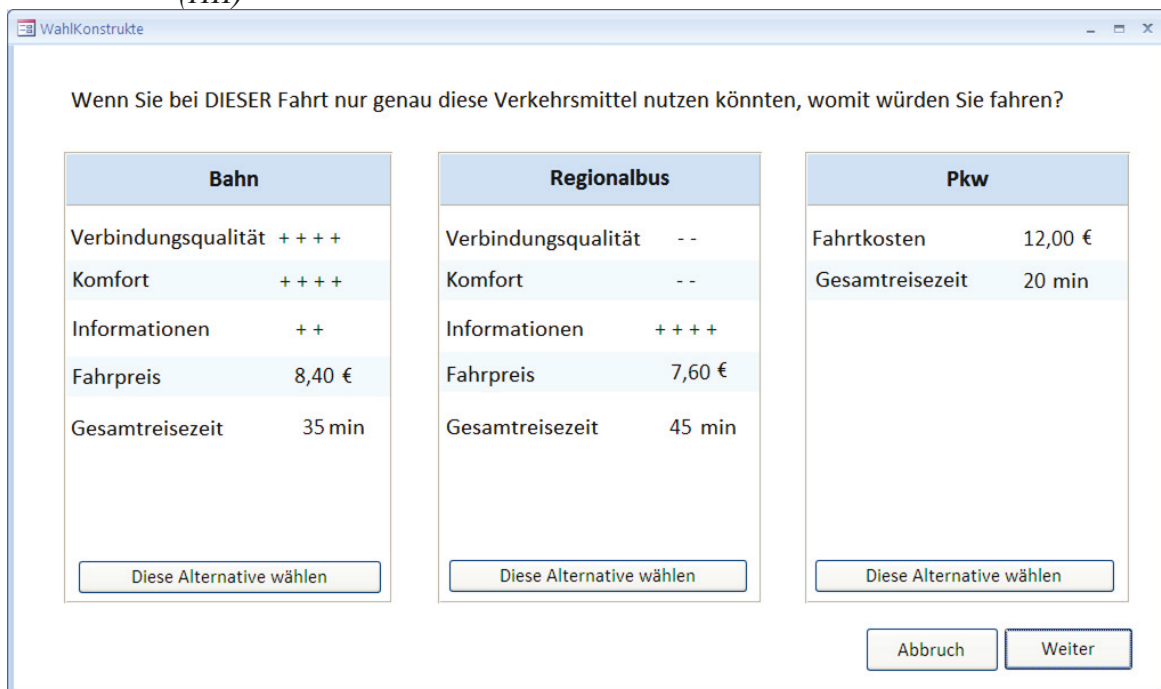
Quelle: Eigene Darstellung.

Abb. D.9: Screenshot des vierten Fragebogenteils mit bimodalem Subexperiment (HII-I) der Informationen



Quelle: Eigene Darstellung.

Abb. D.10: Screenshot des fünften Fragebogenteils mit trimodalem Brückenexperiment (HII)



Quelle: Eigene Darstellung.

Abb. D.11: Screenshot des sechsten Fragebogenteils mit soziodemografischen Fragen und Fragen zum allgemeinen Verkehrsverhalten

The screenshot shows a window titled "Schluss" with the following questions and input fields:

- Wie oft fahren Sie normalerweise mit dem Zug?
- Wie oft nutzen Sie normalerweise den Pkw? (auch als Mitfahrer)
- Besitzen Sie einen Pkw-Führerschein? ja nein
- Steht Ihnen normalerweise ein Pkw zur Verfügung?
- Besitzen Sie einen Zeitfahrchein? (z.B. Monatskarte, Semesterticket, Schwerbehindertenausweis) ja nein
- Erhalten Sie Fahrtkostenerstattung für diese Fahrt? ja nein
- Ihr Geschlecht: männlich weiblich
- Wie alt sind Sie?
- Wie viele Personen leben in Ihrem Haushalt?
- Wie viele dieser Personen sind unter 18 Jahren?
- Wie hoch ist Ihr Bruttohaushaltseinkommen?

An "Abschließen" button is located in the bottom right corner of the window.

Quelle: Eigene Darstellung.

Abb. D.12: Screenshot des siebten Fragebogenteils mit Danksagung

The screenshot shows a window titled "Danke" with the following text:

Vielen Dank für Ihre Teilnahme an dieser Befragung!

An "Abschließen" button is located in the bottom right corner of the window.

Quelle: Eigene Darstellung.

Abb. D.13: Screenshot des achten Fragebogenteils für Anmerkungen des Interviewers

The screenshot shows a window titled "Anmerkungen" with a standard Windows-style title bar. Inside the window, the text "Anmerkungen des Interviewers" is centered above a large, empty rectangular text input field. Below the input field, the text "(max. 255 Zeichen)" is displayed. In the bottom right corner of the window, there is a button labeled "Neuer Datensatz".

Quelle: Eigene Darstellung.

E. Erhebungsinstrument mit HII-I für Pkw-Nutzer

Abb. E.1: Screenshot des ersten Fragebogenteils mit Anmeldung des Interviewers

The screenshot shows a web application window titled 'Start_Export'. The header includes the logo of Fachhochschule Gelsenkirchen, the title 'QUASIMIR - Qualitätssimulationsmodell im Regionalverkehr Hauptuntersuchung', and the text 'Version 2.1'. On the right, it says 'GEFÖRDERT VOM Bundesministerium für Bildung und Forschung'. The main content area contains a form with the following elements:

- 'Interviewernummer:' followed by a text input field.
- 'Erhebungsort' followed by a dropdown menu.
- A button labeled 'Zur Datenerhebung'.
- An 'Exportieren' button.
- A note: '(speichert die Dateien im Ordner C:\Users\{fssv}\Documents\QuasimirExporte\ab)'.

Quelle: Eigene Darstellung.

Abb. E.2: Screenshot des zweiten Fragebogenteils mit Fragen zur Zielgruppenauswahl

The screenshot shows a web application window titled 'Zielgruppenauswahl'. The header includes the text 'Schönen guten Tag!' and 'Wir führen eine Studie zum Verkehrsmittelwahlverhalten durch. Ihre Teilnahme an der Befragung dauert ca. 15 Minuten.' The main content area contains a series of questions with corresponding input fields:

- 'Besitzen Sie einen Pkw-Führerschein?' with radio buttons for 'ja' and 'nein'.
- 'Steht Ihnen normalerweise ein Pkw zur Verfügung?' with a dropdown menu.
- 'Sind Sie berufstätig?' with radio buttons for 'ja' and 'nein'.
- 'Wie erreichen Sie Ihre Arbeitsstätte?' with a dropdown menu.
- 'Wie lange dauert Ihre Fahrt zur Arbeitsstätte insgesamt?' with a text input field followed by 'Minuten'.
- 'Führen Sie regelmäßig Fahrten mit dem Pkw durch, bei denen die Fahrtzeit (eine Richtung) zwischen 20 und 90 Minuten liegt?' with radio buttons for 'ja' and 'nein'.

A 'Weiter' button is located at the bottom right of the form.

Quelle: Eigene Darstellung.

Abb. E.3: Screenshot des dritten Fragebogenteils zur Beschreibung einer konkreten Fahrt

Bitte beschreiben Sie Ihre Fahrt zur Arbeitsstätte!

Zu welchem Zweck führen Sie diese Fahrt durch? von nach

An welchem Ort beginnt diese Fahrt?

An welchem Ort endet diese Fahrt?

Wie lange dauert diese Fahrt mit dem Pkw insgesamt? Minuten (Angaben jeweils für eine Richtung)

Wie hoch sind Ihre Fahrtkosten für diese Fahrt? €

Wie hoch andere Kosten für diese Fahrt? (z.B. Parkkosten) €

Könnten Sie diese Fahrt auch mit der Bahn durchführen? ja nein

An welchem Bahnhof würde diese Fahrt beginnen?

Wie würden Sie diesen Bahnhof erreichen?

An welchem Bahnhof würde für Sie diese Fahrt enden?

Wie würden Sie von dort aus Ihr Ziel erreichen?

Wie oft müssen Sie bei dieser Fahrt zwischen öffentlichen Verkehrsmitteln umsteigen?

Wie lange würde diese Fahrt mit der Bahn insgesamt dauern? Minuten

Was würde Sie diese Fahrt insgesamt mit der Bahn kosten? €

Quelle: Eigene Darstellung.

Abb. E.4: Screenshot des vierten Fragebogenteils mit Rating-Experiment der Verbindungsqualität

Bitte bewerten Sie die Verbindungsqualität des folgenden fiktiven Bahnangebotes!

Pünktlichkeit pünktlich bzw. max. 3 Minuten Verspätung 3 bis 10 Minuten Verspätung 10 bis 20 Minuten Verspätung

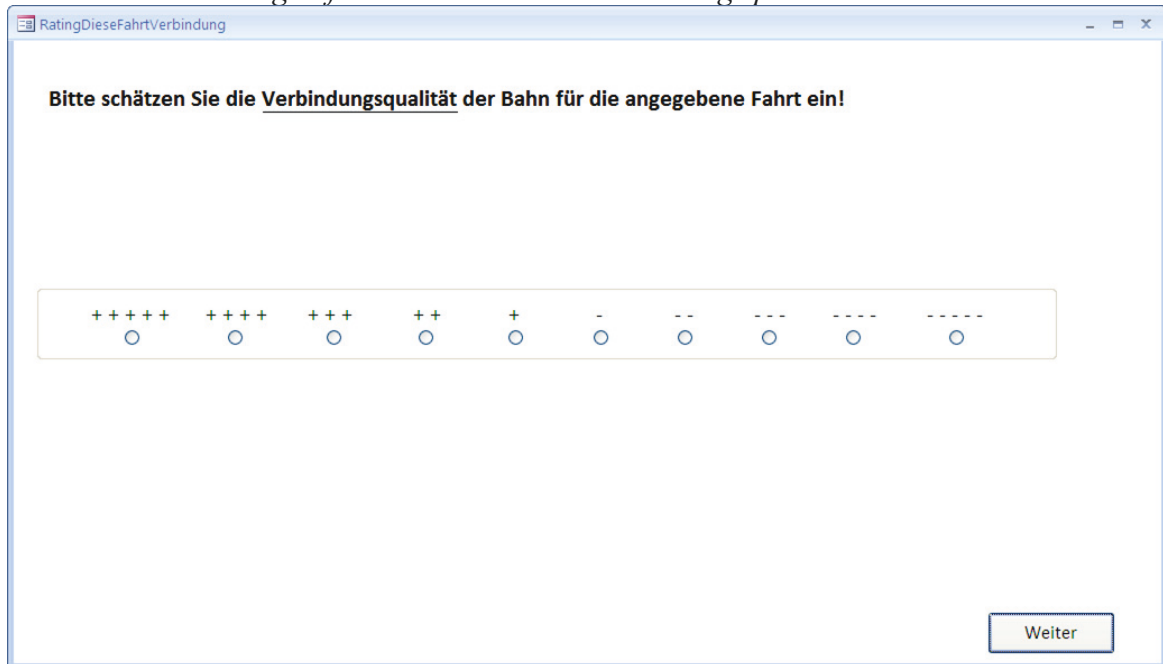
Umsteigehäufigkeit 0 x Umsteigen 1 x Umsteigen mit Anschlussgarantie 1 x Umsteigen ohne Anschlussgarantie

Takt alle 30 Minuten alle 60 Minuten alle 120 Minuten

Bewertung +++++ ++++ +++ ++ + - -- --- ---- -----

Quelle: Eigene Darstellung.

Abb. E.5: Screenshot des vierten Fragebogenteils mit Bewertung der angegebenen Fahrt in Bezug auf das Konstrukt der Verbindungsqualität



Quelle: Eigene Darstellung.

Abb. E.6: Screenshot des fünften Fragebogenteils mit trimodalem Subexperiment (HII-I) der Verbindungsqualität



Quelle: Eigene Darstellung.

Abb. E.7: Screenshot des sechsten Fragebogenteils mit trimodalem Brückenexperiment (HII)

Wenn Sie bei DIESER Fahrt nur genau diese Verkehrsmittel nutzen könnten, womit würden Sie fahren?

| Bahn | | Regionalbus | | Pkw | |
|---------------------|--------|---------------------|--------|-----------------|--------|
| Verbindungsqualität | ++++ | Verbindungsqualität | ++ | Fahrtkosten | 8,80 € |
| Komfort | ++ | Komfort | -- | Gesamtreisezeit | 20 min |
| Informationen | ++ | Informationen | ++++ | | |
| Fahrpreis | 5,10 € | Fahrpreis | 4,60 € | | |
| Gesamtreisezeit | 40 min | Gesamtreisezeit | 30 min | | |

Buttons: Diese Alternative wählen (für jede Alternative), Weiter

Quelle: Eigene Darstellung.

Abb. E.8: Screenshot des siebten Fragebogenteils mit soziodemografischen Fragen und Fragen zum allgemeinen Verkehrsverhalten

Wie oft fahren Sie normalerweise mit dem Zug?

Wie oft nutzen Sie normalerweise den Pkw?
(auch als Mitfahrer)

Besitzen Sie einen Zeitfahrchein? ja nein
(z.B. Monatskarte, Semesterticket, Schwerbehindertenausweis)

Erhalten Sie Fahrtkostenerstattung für die angegebene Fahrt? ja nein

Ihr Geschlecht: männlich weiblich

Wie alt sind Sie?

Wie viele Personen leben in Ihrem Haushalt?

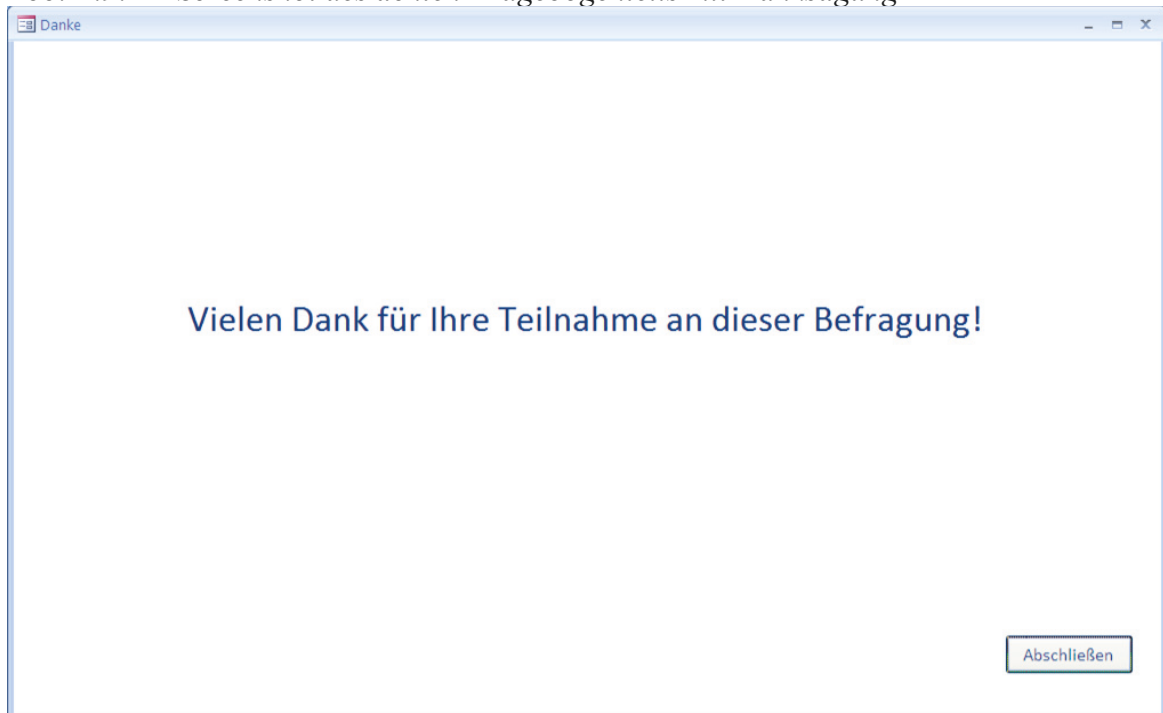
Wie viele dieser Personen sind unter 18 Jahren?

Wie hoch ist Ihr Bruttohaushaltseinkommen?

Button: Abschließen

Quelle: Eigene Darstellung.

Abb. E.9: Screenshot des achten Fragebogenteils mit Danksagung



Quelle: Eigene Darstellung.

Abb. E.10: Screenshot des neunten Fragebogenteils für Anmerkungen des Interviewers



Quelle: Eigene Darstellung.

Literaturverzeichnis

- Adamowicz, W., J. Louviere, M. Williams* (1994): Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities, in: *Journal of Environmental Economics and Management*, Vol. 26, No. 3, S. 271–292. doi:10.1006/jeem.1994.1017.
- Albaum, G.* (1989): BRIDGER, Version 1.0; SIMGRAF, Version 1.0 (Review), in: *Journal of Marketing Research*, Vol. 26, No. 4, S. 486–488, <http://www.jstor.org/stable/3172773> [Zuletzt geprüft am: 17.04.2013].
- Allenby, G.M., J.L. Ginter* (1995): Using Extremes to Design Products and Segment Markets, in: *Journal of Marketing Research*, Vol. 32, No. 4, S. 392–403. doi:10.2307/3152175.
- Anderson, N.H.* (1981): *Foundations of information integration theory*, New York/NY.
- Anderson, N.H.* (1982): *Methods of information integration theory*, New York/NY.
- Angermüller, W.M.* (2007): Die Kundenzufriedenheit und ihr Einfluss auf die Nachfrage: Ein strategisches Instrument zur Erlössteuerung, in: *Der Nahverkehr*, Jg. 25, Nr. 7-8, S. 56–61.
- Arbeitskreis „Fahrzeuge“ der BAG-SPNV* (2006): Empfehlungen für Anforderungen an Fahrzeuge in Vergabeverfahren für Mitglieder der BAG-SPNV, Berlin, <http://spnv.de/website/cms/upload/fakten/AnforderungenFahrzeuge.pdf> [Zuletzt geprüft am: 16.04.2013].
- Arbeitskreis „Fahrzeuge“ der BAG-SPNV* (2010): Empfehlungen für Anforderungen an Fahrzeuge in Vergabeverfahren für Mitglieder der BAG-SPNV, 2. Ausgabe, Berlin, http://spnv.de/website/cms/upload/fakten/2010-07-14_Anforderungen_an_Fahrzeuge_in_Vergabeverfahren_Endversion.pdf [Zuletzt geprüft am: 16.04.2013].
- Argiolu, R. et al.* (2008): Intelligent transport systems and preferences for office locations, in: *Environment and Planning A*, Vol. 40, No. 7, S. 1744–1759. doi:10.1068/a39231.
- Bacher, J.* (1996): *Clusteranalyse: Anwendungsorientierte Einführung*, 2. Aufl., München u.a.
- Backhaus, K., B. Erichson, R. Weiber* (2011): *Fortgeschrittene multivariate Analysemethoden: Eine anwendungsorientierte Einführung*, Berlin u.a.
- Bastians* (2009): Preiselastizitäten im öffentlichen Personenverkehr (ÖPV): Anwendungspotenziale und ihre Übertragbarkeit im räumlichen Kontext, Kiel, <http://d-nb.info/1019867736/34> [Zuletzt geprüft am: 16.04.2013].
- Bäumer/Hussinger/Pfeiffer* (2010): Bedeutung von Kundenzufriedenheit im ÖPNV: Kundenzufriedenheitsmodelle und Messverfahren als Grundlage für die Ableitung von Maßnahmen der Angebotsgestaltung im Marketing-Mix, Schlussbericht FOPS-Projekt FE 70.796/2006, Mannheim.

- Bäumer, M. et al.* (2009): Kundenzufriedenheit und wirtschaftlicher Erfolg im Nahverkehr: Kundenzufriedenheitsmodelle und Messverfahren als Grundlage für die Ableitung von Maßnahmen der Angebotsgestaltung im Marketing-Mix, in: *Der Nahverkehr*, Jg. 27, Nr. 11, S. 22–27.
- Bäumer, M., M. Pfeiffer, W. Meier* (2012): Kundenzufriedenheit: Indikator für unternehmerischen Erfolg?: Ergebnisse einer wissenschaftlichen Untersuchung zum Zusammenhang von Kundenzufriedenheit und ökonomischem Erfolg, in: *Der Nahverkehr*, Jg. 30, Nr. 4, S. 7–13.
- Bäumer/Pfeiffer* (2011): Leitfaden für die Durchführung von Kundenzufriedenheitsbefragungen im ÖPNV, Berlin und Mannheim, http://mitglieder.vdv.de/module/layout_upload/leitfaden_kundenzufriedenheitserhebungen.pdf [Zuletzt geprüft am: 16.04.2013].
- Bech, M., D. Gyrd-Hansen* (2005): Effects coding in discrete choice experiments, in: *Health Economics*, Vol. 14, No. 10, S. 1079–1083. doi:10.1002/hec.984.
- Ben-Akiva, M.E., S.R. Lerman* (1987): *Discrete Choice Analysis: Theory and application to travel demand*, MIT Press Series in transportation studies, 2. Aufl., Cambridge/MA.
- Berekoven, L., W. Eckert, P. Ellenrieder* (2004): *Marktforschung: Methodische Grundlagen und praktische Anwendung*, 10. Aufl., Wiesbaden.
- Beutin, N.* (2006): Verfahren zur Messung der Kundenzufriedenheit im Überblick, in: Homburg, C. (Hrsg.): *Kundenzufriedenheit: Konzepte - Methoden - Erfahrungen*, Wiesbaden, S. 121–169.
- Bijmolt, T.H., M. Wedel* (1995): The effects of alternative methods of collecting similarity data for Multidimensional Scaling, in: *International Journal of Research in Marketing*, Vol. 12, No. 4, S. 363–371. doi:10.1016/0167-8116(95)00012-7.
- Boltze et al.* (2002): Grundlagen für die Beeinflussung des individuellen Verkehrsmittelwahlverhaltens durch Direktmarketing: Schlussbericht. Schriftenreihe des Fachgebiets Verkehrsplanung und Verkehrstechnik, Nr. 9, Darmstadt, http://www.verkehr.tu-darmstadt.de/media/verkehr/fgvv/prof_boltze/BoVeroeff53.pdf [Zuletzt geprüft am: 16.04.2013].
- Borg* (2000): Explorative Multidimensionale Skalierung. ZUMA How-to-Reihen, Nr. 1, Mannheim, http://gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/howto/how-to1ib.pdf [Zuletzt geprüft am: 16.04.2013].
- Borg, I.* (2010): Multidimensionale Skalierung, in: Wolf, C., Best, H. (Hrsg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*, Wiesbaden, S. 391–418.
- Borg, I., P.J.F. Groenen* (2005): *Modern multidimensional scaling: Theory and applications*, New York/NY.
- Borg, I., T. Staufenbiel* (1997): *Theorien und Methoden der Skalierung: Eine Einführung*, 3. Aufl., Bern.
- Borg, I., T. Staufenbiel* (2007): *Lehrbuch: Theorien und Methoden der Skalierung*, 4. Aufl., Bern.

- Bortz, J.* (2005): Statistik für Human- und Sozialwissenschaftler, 6. Aufl., Berlin u.a.
- Bortz, J., N. Döring* (2006): Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler, 4. Aufl., Berlin u.a.
- Bortz, J., G.A. Lienert* (2008): Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben, 3. Aufl., Berlin u.a.
- Bos, I.D.M. et al.* (2003): Cognition and Relative Importance Underlying Consumer Valuation of Park-and-Ride Facilities, in: Transportation Research Record, Vol. 1835, No. 1, S. 121–127. doi:10.3141/1835-15.
- Bos, I.D.M.* (2004): Changing seats: A behavioural analysis of P&R use, Delft, Niederlande.
- Bos, I.D.M. et al.* (2004): The choice of park and ride facilities: an analysis using a context-dependent hierarchical choice experiment, in: Environment and Planning A, Vol. 36, No. 9, S. 1673–1686. doi:10.1068/a36138.
- Bos, I.D.M., E.J.E. Molin, R.E.C.M. van der Heijden* (2002): User-based definition of higher-order decision constructs to model preferences for P&R transfer points, in: Sucharov, L.J. (Hrsg.): Urban transport VIII: Urban transport and the environment in the 21st century, Southampton/UK u.a., S. 417–426.
- Bos, I., E. Molin* (2006): Is there a 'Stick' Bonus? A Stated Choice Model for P&R Patronage incorporating Cross-Effects, in: European Journal of Transport and Infrastructure Research, Vol. 6, No. 3, S. 275–290, http://www.ejtir.tbm.tudelft.nl/issues/2006_03/pdf/2006_03_04.pdf [Zuletzt geprüft am: 16.04.2013].
- Bracher, T. et al.* (2004): ÖPNV im Wettbewerb: Management-Planspiel in der Region Berlin, Berlin.
- Bradley, M.A., A.J. Daly* (1997): Estimation of Logit Choice Models Using Mixed Stated-Preference and Revealed-Preference Information, in: Stopher, P.R., Lee-Gosselin, M. (Hrsg.): Understanding travel behaviour in an era of change, Oxford/UK, S. 209–232.
- Bradley, M., A. Daly* (1992): Uses of the Logit Scaling Approach in Stated Preference Analysis, in: World Conference on Transport Research Society (Hrsg.): Traffic, Demand and Network Modelling, Proceedings of the 6th World Conference on Transport Research 2, Lyon, S. 811–823.
- Bradley, M., A. Daly* (1994): Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data, in: Transportation, Vol. 21, No. 2, S. 167–184. doi:10.1007/BF01098791.
- Bronstein, I.* (1999): Taschenbuch der Mathematik, 4. Aufl., Frankfurt am Main.
- Bühl, A.* (2009): PASW 18: Einführung in die moderne Datenanalyse, 12. Aufl., München u.a.
- Bühner, M.* (2006): Einführung in die Test- und Fragebogenkonstruktion, 2. Aufl., München.

- Bundesministerium für Verkehr, Bau und Stadtentwicklung* (2011): Verantwortungstragen - Zukunftsgestalten: Nachhaltigkeitsbericht des Bundesministeriums für Verkehr, Bau und Stadtentwicklung, Berlin.
- Büschken, J., C. von Thaden* (2000): Clusteranalyse, in: Herrmann, A., Homburg, C. (Hrsg.): Marktforschung: Methoden, Anwendungen, Praxisbeispiele, Gabler-Lehrbuch, Wiesbaden, S. 337–380.
- Caussade, S. et al.* (2005): Assessing the influence of design dimensions on stated choice experiment estimates, in: *Transportation Research Part B: Methodological*, Vol. 39, No. 7, S. 621–640. doi:10.1016/j.trb.2004.07.006.
- Chiang, L.-S., J.-L. Lu, H.-C. Chang* (2003): Modeling the Effect of Destination Attributes on the Intercity Travelers' Mode Choice Behavior in Taiwan Area, in: *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 4, S. 717–730, www.easts.info/2003proceedings/papers/0717.pdf [Zuletzt geprüft am: 16.04.2013].
- Coles, S.* (2004): *An introduction to statistical modeling of extreme values*, 3. Aufl., London/UK u.a.
- Commandeur/Heiser* (1993): Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices, No. RR-93-04, http://www.datatheory.nl/pdfs/93/93_04.pdf [Zuletzt geprüft am: 18.01.2012].
- Coxon, A.P.M.* (1999): *Sorting data: Collection and analysis*, Sage University Papers on Quantitative Applications in the Social Sciences, No. 07-127, Thousand Oaks/CA u.a.
- Cox, T.F., M.A.A. Cox* (2001): *Multidimensional scaling*, 2. Aufl., Boca Raton/FL.
- Dillman, D.A. et al.* (2002): Survey Nonresponse in Design, Data Collection, and Analysis, in: Groves, R.M. et al. (Hrsg.): *Survey nonresponse*, New York/NY u.a., S. 3–26.
- Eichmann, V. et al.* (2006): *Umweltfreundlicher, attraktiver und leistungsfähiger ÖPNV - ein Handbuch*, Berlin.
- Engehausen, R.* (1999): Bausteine für eine hohe Kundenzufriedenheit, in: Töpfer, A. (Hrsg.): *Kundenzufriedenheit messen und steigern*, Neuwied, S. 197–223.
- Fahrmeir, L., Hamerle, A., Tutz, G.* (Hrsg.) (1996): *Multivariate statistische Verfahren*, 2. Aufl., Berlin u.a.
- Felscher-Suhr, U., H. Meyer, W. Legath* (2011): Qualitätsvergleich mit dem ÖPNV-Qualitätsbarometer: Ein Benchmarking auf Basis ermittelter Qualitätskennwerte, in: *Der Nahverkehr*, Jg. 29, Nr. 9, S. 41–47.
- Fiebig, D.G. et al.* (2010): The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity, in: *Marketing Science*, Vol. 29, No. 3, S. 393–421. doi:10.1287/mksc.1090.0508.
- Fischer, J.* (2001): *Individualisierte Präferenzanalyse: Entwicklung und empirische Prüfung einer vollkommen individualisierten Conjoint-Analyse*, Wiesbaden.

- Forschungsgesellschaft für Straßen- und Verkehrswesen* (2006): Hinweise für die Qualitätssicherung im ÖPNV, Köln.
- Georgi, A., C. Wüst-Rocktäschel, J. Heller* (2006): Was Fahrgäste wirklich wollen: Die Bevölkerungsbefragung der Offenbacher Verkehrs-Betriebe, in: *Der Nahverkehr*, Jg. 24, Nr. 3, S. 57–61.
- Giering, A.* (2000): Der Zusammenhang zwischen Kundenzufriedenheit und Kundenloyalität: Eine Untersuchung moderierender Effekte, Schriftenreihe des Instituts für Marktorientierte Unternehmensführung, Wiesbaden.
- Gläser, J., G. Laudel* (2009): Experteninterviews und qualitative Inhaltsanalyse: Als Instrumente rekonstruierender Untersuchungen, 3. Aufl., Wiesbaden.
- Gorter et al.* (2000): Bedeutung "weicher" Angebotsmerkmale bei der Angebotsgestaltung sowie ihre Berücksichtigung bei der Einnahmenaufteilung und Vergabe von Leistungen im Wettbewerb: Leitfaden zur Vereinbarung von Qualitätsstandards in Verträgen, im Auftrag des Bundesministers für Verkehr, Forschungsbericht FE-Nr. 70.543/1998/, Berlin.
- Gould, W., J.S. Pitblado, B. Poi* (2010): Maximum likelihood estimation with Stata, 4. Aufl., College Station/TX.
- Green, P.E., F.J. Carmone, S.M. Smith* (1989): Multidimensional scaling: Concepts and applications, Boston/MA.
- Green, P.E., A.M. Krieger, M.K. Agarwal* (1993): A cross validation test of four models for quantifying multiattribute preferences, in: *Marketing Letters*, Vol. 4, No. 4, S. 369–380. doi:10.1007/BF00994355.
- Green, P.E.S.V.* (1978): Conjoint Analysis in Consumer Research: Issues and Outlook, in: *Journal of Consumer Research*, Vol. 5, No. 2, S. 103–123.
- Green, P.E., V. Srinivasan* (1990): Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice, in: *Journal of Marketing*, Vol. 54, No. 4, S. 3–19, <http://www.jstor.org/stable/1251756> [Zuletzt geprüft am: 16.04.2013].
- Groves, R.M. et al.* (Hrsg.) (2002): Survey nonresponse, New York/NY u.a.
- Hamerle, A., H. Pape* (1996): Grundlagen der mehrdimensionalen Skalierung, in: *Fahrmeir, L., Hamerle, A., Tutz, G.* (Hrsg.): *Multivariate statistische Verfahren*, Berlin u.a., S. 767–793.
- Hartung, J., B. Elpelt* (1995): *Multivariate Statistik: Lehr- und Handbuch der angewandten Statistik*, 5. Aufl., München.
- Hausman, J., D. McFadden* (1984): Specification Tests for the Multinomial Logit Model, in: *Econometrica*, Vol. 52, No. 5, S. 1219–1240.
- Helm, R., M. Steiner* (2008): *Präferenzmessung: Methodengestützte Entwicklung zielgruppenspezifischer Produktinnovationen*, Stuttgart.
- Helten/Fischer* (2004): Urbaneye Working Paper No. 13: What do people think about CCTV? Findings from a Berlin survey, Berlin, www.urbaneye.net/results/ue_wp13.pdf [Zuletzt geprüft am: 16.04.1013].

- Hempel/Töpfer* (2004): Urbaneye Working Paper 15: CCTV in Europe. Final Report, Berlin, www.urbaneye.net/results/ue_wp15.pdf [Zuletzt geprüft am: 16.04.2013].
- Hensher, D.A.* (1991): Hierarchical Stated Response Designs - An Application to Bus User Preferences, in: *Logistics and Transportation Review*, Vol. 26, No. 4, S. 299–321.
- Hensher, D.A.* (2004): Identifying the Influence of Stated Choice Design Dimensionality on Willingness to Pay for Travel Time Savings, in: *Journal of Transport Economics and Policy*, Vol. 38, No. 3, S. 425–446. doi:10.2307/20173065.
- Hensher, D.A.* (2008): Empirical approaches to combining revealed and stated preference data: Some recent developments with reference to urban mode choice, in: *Research in Transportation Economics*, Vol. 23, No. 1, S. 23–29. doi:10.1016/j.retrec.2008.10.003.
- Hensher, D.A., M. Bradley* (1993): Using stated response choice data to enrich revealed preference discrete choice models, in: *Marketing Letters*, Vol. 4, No. 2, S. 139–151. doi:10.1007/BF00994072.
- Hensher, D.A., W.H. Greene* (2002): Specification and estimation of the nested logit model: alternative normalisations, in: *Transportation Research Part B: Methodological*, Vol. 36, No. 1, S. 1–17. doi:10.1016/S0191-2615(00)00035-7.
- Hensher, D.A., J.J. Louviere, J.D. Swait* (1998): Combining sources of preference data, in: *Journal of Econometrics*, Vol. 89, No. 1-2, S. 197–221. doi:10.1016/S0304-4076(98)00061-X.
- Hensher, D.A., P. Prioni* (2002): A Service Quality Index for Area-wide Contract Performance Assessment, in: *Journal of Transport Economics and Policy*, Vol. 36, No. 1, S. 93–113, <http://www.jstor.org/stable/20053894> [Zuletzt geprüft am: 17.04.2013].
- Hensher, D.A., J.M. Rose, W.H. Greene* (2005): *Applied choice analysis*, Cambridge/UK u.a.
- Hensher, D.A., J.M. Rose, W.H. Greene* (2008): Combining RP and SP data: biases in using the nested logit ‘trick’ – contrasts with flexible mixed logit incorporating panel and scale effects, in: *Journal of Transport Geography*, Vol. 16, No. 2, S. 126–133. doi:10.1016/j.jtrangeo.2007.07.001.
- Hensher, D.A., P. Stopher, P. Bullock* (2003): Service quality - developing a service quality index in the provision of commercial bus contracts, in: *Transportation Research Part A: Policy and Practice*, Vol. 37, No. 6, S. 499–517. doi:10.1016/S0965-8564(02)00075-7.
- Hensher, D., Stopher, P.* (Hrsg.) (1979): *Behavioural Travel Modelling*, London/UK.
- Herrmann, A., Homburg, C.* (Hrsg.) (2000): *Marktforschung: Methoden, Anwendungen, Praxisbeispiele*, Gabler-Lehrbuch, 2. Aufl., Wiesbaden.
- Hinterhuber, H.H., Matzler, K.* (Hrsg.) (2002): *Kundenorientierte Unternehmensführung: Kundenorientierung, Kundenzufriedenheit, Kundenbindung*, 3. Aufl., Wiesbaden.

- Holz-Rau et al.* (2010): ÖPNV: Planung für ältere Menschen.: Ein Leitfaden für die Praxis., BMVBS-Online-Publikation 09/2010, Berlin, <http://d-nb.info/1003341470/34> [Zuletzt geprüft am: 16.04.2013].
- Homburg, C.* (Hrsg.) (2006): Kundenzufriedenheit: Konzepte - Methoden - Erfahrungen, 6. Aufl., Wiesbaden.
- Homburg, C., R. Stock-Homburg* (2006): Theoretische Perspektiven zur Kundenzufriedenheit, in: Homburg, C. (Hrsg.): Kundenzufriedenheit: Konzepte - Methoden - Erfahrungen, Wiesbaden, S. 17–51.
- Hubert, L., P. Arabie* (1985): Comparing partitions, in: *Journal of Classification*, Vol. 2, No. 1, S. 193–218. doi:10.1007/BF01908075.
- Hunt, G.L.* (2000): Alternative Nested Logit Model Structures and the Special Case of Partial Degeneracy, in: *Journal of Regional Science*, Vol. 40, No. 1, S. 89–113. doi:10.1111/0022-4146.00166.
- IBM Corporation* (2011a): IBM SPSS Statistics 20 Algorithms, http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf [Zuletzt geprüft am: 15.04.2013].
- IBM Corporation* (2011b): IBM SPSS Statistics 20 Command Syntax Reference, http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Command_Syntax_Reference.pdf [Zuletzt geprüft am: 15.04.2013].
- IBM Corporation* (2011c): IBM SPSS Statistics Base 20, http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Base.pdf [Zuletzt geprüft am: 15.04.2013].
- Isfort, A.* (2000): Das ÖPNV-Kundenbarometer: Ein Instrument zur Standortbestimmung für kleine und mittlere Unternehmen und Verbände, in: *Der Nahverkehr*, Jg. 18, Nr. 11, S. 39–42.
- Jankisz, E., H. Moosbrugger* (2007): Planung und Entwicklung von psychologischen Tests und Fragebogen, in: Moosbrugger, H., Kelava, A. (Hrsg.): *Testtheorie und Fragebogenkonstruktion*, Heidelberg, S. 27–72.
- Jann, B.* (2005): Making regression tables from stored estimates, in: *The Stata Journal*, Vol. 5, No. 3, S. 288–308, <http://www.stata-journal.com/sjpdf.html?articlenum=st0085> [Zuletzt geprüft am: 17.04.2013].
- Jann, B.* (2007): Making regression tables simplified, in: *The Stata Journal*, Vol. 7, No. 2, S. 227–244, http://www.stata-journal.com/sjpdf.html?articlenum=st0085_1 [Zuletzt geprüft am: 17.04.2013].
- Johnson, R.M.* (1974): Trade-off Analysis of Consumer Values, in: *Journal of Marketing Research*, Vol. 11, No. 2, S. 121–127, <http://www.jstor.org/stable/3150548> [Zuletzt geprüft am: 17.04.2013].
- Jones, R.A., R.D. Ashmore* (1973): The structure of intergroup perception: Categories and dimensions in views of ethnic groups and adjectives used in stereotype research, in: *Journal of Personality and Social Psychology*, Vol. 25, No. 3, S. 428–438. doi:10.1037/h0034162.

- Kaiser, M.-O.* (2005): Erfolgsfaktor Kundenzufriedenheit: Dimensionen und Messmöglichkeiten, 2. Aufl., Berlin.
- Karopka, H.-J. et al.* (2000): Wie erlebt der Kunde den öffentlichen Nahverkehr?: Qualitativ-psychologische Grundlagenstudie zur ÖPNV-Nutzung, in: *Der Nahverkehr*, Jg. 18, Nr. 11, S. 18–22.
- Kaufmann, H., H. Pape* (1996): Clusteranalyse, in: Fahrmeir, L., Hamerle, A., Tutz, G. (Hrsg.): *Multivariate statistische Verfahren*, Berlin u.a., S. 437–536.
- Kenett, R.S., Salini, S.* (Hrsg.) (2012): *Modern analysis of customer surveys: With applications using R*, Chichester/UK.
- Kmenta, J.* (1997): *Elements of econometrics*, 2. Aufl., Ann Arbor/MI.
- Koschate, N.* (2002): *Kundenzufriedenheit und Preisverhalten: Theoretische und empirisch experimentelle Analysen*, Schriftenreihe des Instituts für Markt-orientierte Unternehmensführung, Wiesbaden.
- Krietemeyer, H., K. Wergles* (2006): Kundenzufriedenheit aktuell messen: Das MVV-Kundenbarometer-Tracking: Hintergründe und Erfahrungen, in: *Der Nahverkehr*, Jg. 24, No. 7-8, S. 28–34.
- Kroeber-Riel, W., P. Weinberg, A. Gröppel-Klein* (2009): *Konsumentenverhalten*, 9. Aufl., München.
- Kruskal, J.B.* (1964a): Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, in: *Psychometrika*, Vol. 29, No. 1, S. 1–27. doi:10.1007/BF02289565.
- Kruskal, J.B.* (1964b): Nonmetric multidimensional scaling: A numerical method, in: *Psychometrika*, Vol. 29, No. 2, S. 115–129. doi:10.1007/BF02289694.
- Kruskal, J.B., M. Wish* (1978): *Multidimensional scaling*, Sage University Papers on Quantitative Applications in the Social Sciences, No. 07-11, Beverly Hills/CA.
- Kurniawan, S.H., P. Zaphiris* (2003): Web Health Information Architecture for Older Users, in: *IT&Society*, Vol. 1, No. 3, S. 42–63, http://www soi.city.ac.uk/~zaphiri/Papers/it_and_s.pdf [Zuletzt geprüft am: 16.04.2013].
- Laaser* (1994): Die Bahnstrukturreform: Richtige Weichenstellung oder Fahrt aufs Abstellgleis?, Working Paper. Kieler Diskussionsbeiträge, Nr. 239, Kiel, <https://www.econstor.eu/dspace/bitstream/10419/821/1/166112445.PDF> [Zuletzt geprüft am: 13.04.2013].
- Lancaster, K.J.* (1966): A New Approach to Consumer Theory, in: *The Journal of Political Economy*, Vol. 74, No. 2, S. 132–157, <http://www.jstor.org/stable/1828835> [Zuletzt geprüft am: 17.04.2013].
- Lawerentz, R., A. Balling, W. Fleischer* (2002): Qualitätsziele müssen messbar sein: Vorgehensweise und Ergebnisse des Unternehmensbereiches Bus der BVG, in: *Der Nahverkehr*, Vol. 20, No. 11, S. 13–19.
- Lawless, H.T., N. Sheng, S.S. Knoops* (1995): Multidimensional scaling of sorting data applied to cheese perception, in: *Food Quality and Preference*, Vol. 6, No. 2, S. 91–98. doi:10.1016/0950-3293(95)98553-U.

- Lehmbrock, M. et al.* (2006): Funktionale Beschreibung von ÖPNV in Städten, Difu-Beiträge zur Stadtforschung, Nr. 44, Berlin.
- Lenz et al.* (2010): Mobilität in Deutschland 2008: Ergebnisbericht: Struktur - Aufkommen - Emissionen - Trends, Bonn und Berlin, http://www.mobilitaet-in-deutschland.de/pdf/MiD2008_Abschlussbericht_I.pdf [Zuletzt geprüft am: 16.04.2013].
- Litz, H.P.* (2000): Multivariate statistische Methoden und ihre Anwendung in den Wirtschafts- und Sozialwissenschaften, München.
- Louviere, J.J.* (1984): Hierarchical Information Integration: A New Approach for the Design and Analysis of Complex Multiattribute Judgment Problems, in: *Advances in Consumer Research*, Vol. 11, S. 148–155.
- Louviere, J.J. et al.* (2000): Stated choice methods: Analysis and applications, 1. Aufl., 6. Druck, Cambridge/UK u.a.
- Louviere, J.J., T.N. Flynn, R.T. Carson* (2010): Discrete Choice Experiments Are Not Conjoint Analysis, in: *Journal of Choice Modelling*, Vol. 3, No. 3, S. 57–72, <http://www.jocm.org.uk/index.php/JOCM/article/view/124/66> [Zuletzt geprüft am: 16.04.2013].
- Louviere, J.J., G.J. Gaeth* (1987): Decomposing the Determinants of Retail Facility Choice Using the Methods of Hierarchical Information Integration: A Supermarket Illustration, in: *Journal of Retailing*, Vol. 63, No. 1, S. 25–48.
- Louviere, J.J., H.J.P. Timmermans* (1990a): A Review of Recent Advances in Decompositional Preference and Choice Models, in: *Tijdschrift voor Economische en Sociale Geografie*, Vol. 81, No. 3, S. 214–224. doi:10.1111/j.1467-9663.1990.tb00772.x.
- Louviere, J.J., H.J.P. Timmermans* (1990b): Hierarchical Information Integration Applied to Residential Choice Behavior, in: *Geographical Analysis*, Vol. 22, No. 2, S. 127–144. doi:10.1111/j.1538-4632.1990.tb00200.x.
- Louviere, J.J., H.J.P. Timmermans* (1990c): Using hierarchical information integration to model consumer responses to possible planning actions: recreation destination choice illustration, in: *Environment and Planning A*, Vol. 22, No. 3, S. 291–308. doi:10.1068/a220291.
- Louviere, J.J., H.J.P. Timmermans* (1992): Testing the external validity of hierarchical conjoint analysis models of recreational destination choice, in: *Leisure Sciences*, Vol. 14, No. 3, S. 179–194. doi:10.1080/01490409209513167.
- Louviere, J.J., G. Woodworth* (1983): Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach on Aggregate Data, in: *Journal of Marketing Research*, Vol. 20, No. 4, S. 350–367.
- Magidson, J., J.K. Vermunt* (2008): Removing the Scale Factor Confound in Multinomial Logit Choice Models to Obtain Better Estimates of Preference, in: *Sawtooth Software (Hrsg.): Proceedings of the Sawtooth Software*

- Conference: October 2007, S. 139–154, <http://statisticalinnovations.com/technicalsupport/sawtooth2007.pdf> [Zuletzt geprüft am: 17.04.2013].
- Manski, C.F., McFadden, D.* (Hrsg.) (1981): Structural analysis of discrete data with econometric applications, Boston/MA.
- Matzler, K., F. Bailom* (2002): Messung von Kundenzufriedenheit, in: Hinterhuber, H.H., Matzler, K. (Hrsg.): Kundenorientierte Unternehmensführung: Kundenorientierung, Kundenzufriedenheit, Kundenbindung, Wiesbaden, S. 213–244.
- Matzler, K., E. Sauerwein, C. Stark* (2002): Methoden zur Identifikation von Basis-, Leitungs- und Begeisterungsfaktoren, in: Hinterhuber, H.H., Matzler, K. (Hrsg.): Kundenorientierte Unternehmensführung: Kundenorientierung, Kundenzufriedenheit, Kundenbindung, Wiesbaden, S. 265–289.
- McFadden, D.* (1974): Conditional logit analysis of qualitative choice behavior, in: Zarembka, P. (Hrsg.): Frontiers in econometrics, New York/NY, S. 105–142.
- McFadden, D.* (1979): Quantitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments, in: Hensher, D., Stopher, P. (Hrsg.): Behavioural Travel Modelling, London/UK, S. 279–318.
- McFadden, D.* (1981): Econometric Models of Probabilistic Choice, in: Manski, C.F., McFadden, D. (Hrsg.): Structural analysis of discrete data with econometric applications, Boston/MA, S. 198–272.
- Meffert, H.* (Hrsg.) (2000): Verkehrsdienstleistungsmarketing: Marktorientierte Unternehmensführung bei der Deutschen Bahn AG, Wiesbaden.
- Meffert, H., M. Bruhn* (2003): Dienstleistungsmarketing: Grundlagen - Konzepte - Methoden, Mit Fallstudien, Meffert Marketing Edition, 4. Aufl., Wiesbaden.
- Meulman/Heiser* (2011): IBM SPSS Categories 20, Chicago/IL, http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Categories.pdf [Zuletzt geprüft am: 16.04.2013].
- Miller, G.A.* (1969): A psychological method to investigate verbal concepts, in: Journal of Mathematical Psychology, Vol. 6, No. 2, S. 169–191. doi:10.1016/0022-2496(69)90001-7.
- Milligan, G.W., M.C. Cooper* (1986): A Study of the Comparability of External Criteria for Hierarchical Cluster Analysis, in: Multivariate Behavioral Research, Vol. 21, No. 4, S. 441–458. doi:10.1207/s15327906mbr2104_5.
- Molin, E.J.E.* (1999): Conjoint modeling approaches for residential group preferences, Eindhoven.
- Molin, E.J.E., H.J.P. Timmermans* (2003): Testing hierarchical information integration theory: the causal structure of household residential satisfaction, in: Environment and Planning A, Vol. 35, No. 1, S. 43–58. doi:10.1068/a3416.
- Molin, E.J.E., H.J.P. Timmermans* (2009): Hierarchical Information Integration Experiments and Integrated Choice Experiments, in: Transport Reviews, Vol. 29, No. 5, S. 635–655. doi:10.1080/01441640902829470.

- Molin, E.J.E., M. van Gelder* (2008): Freeway Access to Public Transport: A Hierarchical Multimodal Choice Model, in: *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2076, No. 1, S. 106–113. doi:10.3141/2076-12.
- Molin, E.J., H. Oppewal, H.J. Timmermans* (2000): A Comparison of Full Profile and Hierarchical Information Integration Conjoint Methods to Modeling Group Preferences, in: *Marketing Letters*, Vol. 11, No. 2, S. 165–175. doi:10.1023/A:1008194807943.
- Molin, E.J., H. Oppewal, H.J. Timmermans* (2001): Analyzing heterogeneity in conjoint estimates of residential preferences, in: *Journal of Housing and the Built Environment*, Vol. 16, No. 3-4, S. 267–284. doi:10.1023/A:1012539415696.
- Moosbrugger, H., A. Kelava* (2007): Qualitätsanforderungen an einen psychologischen Test (Testgütekriterien), in: *Moosbrugger, H., Kelava, A. (Hrsg.): Testtheorie und Fragebogenkonstruktion*, Heidelberg, S. 7–26.
- Moosbrugger, H., Kelava, A. (Hrsg.)* (2007): *Testtheorie und Fragebogenkonstruktion*, Heidelberg.
- Nahverkehr Westfalen-Lippe* (2011): Nahverkehrsplan Westfalen-Lippe: Erster Nahverkehrsplan für den SPNV im NWL, Münster, <http://www.nwl-info.de/dokumente/pdf/service/broschueren/nahverkehrsplan.pdf> [Zuletzt geprüft am: 16.04.2013].
- Neidell, L.A.* (1972): Procedures for Obtaining Similarities Data, in: *Journal of Marketing Research*, Vol. 9, No. 3, S. 335–337.
- Norojono, O., W. Young* (2003): A Stated preference freight mode choice model, in: *Transportation Planning and Technology*, Vol. 26, No. 2, S. 195–212. doi:10.1080/715020600.
- ÖPNV-Zukunftskommission NRW* (2012): Zwischenbericht über die bisherige Kommissionsarbeit, Düsseldorf, http://www.mbwsv.nrw.de/aktuelle_Meldungen_Startseite/2013_01_11_PNV_Zwischenbericht/Zwischenbericht_Zukunftskommission_2013-01-11.pdf [Zuletzt geprüft am: 14.04.2013].
- Oppewal, H. et al.* (1997): Modelling the effects of shopping centre size and store variety on consumer choice behaviour, in: *Environment and Planning A*, Vol. 29, No. 6, S. 1073–1090. doi:10.1068/a291073.
- Oppewal, H., M. Klabbers* (2003): Compromising Between Information Completeness and Task Simplicity: A Comparison of Self-Explicated, Hierarchical Information Integration, and Full-profile Conjoint Methods, in: *Advances in Consumer Research*, Vol. 30, S. 298–304.
- Oppewal, H., J.J. Louviere, H.J.P. Timmermans* (1994): Modeling Hierarchical Conjoint Processes With Integrated Choice Experiments, in: *Journal of Marketing Research*, Vol. 31, S. 92–105. doi:10.2307/3151949.
- Oppewal, H., M. Vriens* (2000): Measuring perceived service quality using integrated conjoint experiments, in: *International Journal of Bank Marketing*, Vol. 18, No. 4, S. 154–169. doi:10.1108/02652320010349049.

- Panitz, P., K. Jobke, T. Jensch* (2012): Bürgerbeteiligung im Nahverkehr: Mitwirkung der Fahrgäste an der Fahrplanerstellung im SPNV, in: *Der Nahverkehr*, Jg. 30, No. 1-2, S. 60–61.
- Parasuraman, A., V.A. Zeithaml, L.L. Berry* (1985): A Conceptual Model of Service Quality and Its Implications for Future Research, in: *Journal of Marketing*, Vol. 49, No. 4, S. 41–50.
- Perrey, J.* (1998): Nutzenorientierte Marktsegmentierung: Ein integrativer Ansatz zum Zielgruppenmarketing im Verkehrsdienstleistungsbereich, Wiesbaden.
- Perrey, J.* (2000): Nutzenorientierte Marktsegmentierung im Verkehrsdienstleistungsbereich: Ein integrativer Ansatz zum Zielgruppenmarketing, in: Meffert, H. (Hrsg.): *Verkehrsdienstleistungsmarketing: Marktorientierte Unternehmensführung bei der Deutschen Bahn AG*, Wiesbaden, S. 57–118.
- Pfeiffer, M., J. Schmidt* (2006): Auswertung von Fahrgastbefragungen: Punkt- und Intervallschätzung bei komplexen Stichprobendesigns, in: *Der Nahverkehr*, Jg. 24, Nr. 5, S. 66–69.
- Plantsch, M.* (2011): Einfluss kritischer Einflüsse im ÖPNV auf Kundenzufriedenheit: Ergebnisse einer Studie nach der Critical Incident Technique, in: *Der Nahverkehr*, Jg. 29, Nr. 6, S. 19–23.
- Plath, B.* (2009): ÖPNV-Qualität verbundweit steuern: Erste Erfahrungen des Hamburger Verkehrsverbundes mit einem Qualitätssteuerungsverfahren, in: *Der Nahverkehr*, Jg. 27, Nr. 4, S. 47–50.
- Probst, G., C. Kahrs, S. Rieschick-Dziabas* (2001): Was bindet Fahrgäste dauerhaft?: Ereignisorientierte Kundenzufriedenheit im ÖPNV, in: *Der Nahverkehr*, Jg. 19, Nr. 3, S. 8–12.
- Pullman, M.E., K.J. Dodson, W.L. Moore* (1999): A Comparison of Conjoint Methods When There Are Many Attributes, in: *Marketing Letters*, Vol. 10, No. 2, S. 125–138. doi:10.1023/A:1008036829555.
- Ramírez-Hurtado, J.M.* (2010): Measuring Preferences: from Conjoint Analysis to Integrated Conjoint Experiments, in: *Revista de Métodos Cuantitativos para la Economía y la Empresa*, Vol. 9, S. 28–43.
- Rand, W.M.* (1971): Objective Criteria for the Evaluation of Clustering Methods, in: *Journal of the American Statistical Association*, Vol. 66, No. 336, S. 846–850. doi:10.2307/2284239.
- Richter, C., S. Keuchel* (2012a): Deriving Constructs for Hierarchical Information Integration Experiments in Regional Public Transport, in: *Procedia - Social and Behavioral Sciences*, Vol. 54, S. 666–674. doi:10.1016/j.sbspro.2012.09.784.
- Richter, C., S. Keuchel* (2012b): Modelling Mode Choice in Passenger Transport with Integrated Hierarchical Information Integration, in: *Journal of Choice Modelling*, Vol. 5, No. 1, S. 1–21, <http://jocm.org.uk/index.php/JOCM/article/download/165/93> [Zuletzt geprüft am: 16.04.2013].
- Ronning, G.* (1991): *Mikroökonomie*, Berlin u.a.

- Rosenberg, S., Nelson Carnot, P.S. Vivekananthan* (1968): A Multidimensional Approach to the Structure of Personality Impressions, in: *Journal of Personality and Social Psychology*, Vol. 9, No. 4, S. 283–294. doi:10.1037/h0026086.
- Rosenberg, S., M. Park Kim* (1975): The Method of Sorting as a Data-Gathering Procedure in Multivariate Research, in: *Multivariate Behavioral Research*, Vol. 10, No. 4, S. 489–502. doi:10.1207/s15327906mbr1004_7.
- Sawtooth Software* (2007): The ACA/Web v6.0: Technical Paper Series, www.sawtoothsoftware.com/download/techpap/acatech.pdf [Zuletzt geprüft am: 26.10.2012].
- Sawtooth Software* (Hrsg.) (2008): *Proceedings of the Sawtooth Software Conference: October 2007*.
- Scarpa, R. et al.* (2012): Inferred and Stated Attribute Non-attendance in Food Choice Experiments, in: *American Journal of Agricultural Economics*, Vol. 95, No. 1, S. 165–180. doi:10.1093/ajae/aas073.
- Schäfer-Sparenberg/Bongardt/Dalkmann* (2006): Steuerung nachhaltiger Daseinsvorsorge im öffentliche Nahverkehr: Politische Handlungsperspektiven. *Wuppertal Papers*, Nr. 161, Wuppertal, <http://epub.wuppertal.org/files/2396/WP161.pdf> [Zuletzt geprüft am: 16.04.2013].
- Scheffler, H.* (2000): Stichprobenbildung und Datenerhebung, in: Herrmann, A., Homburg, C. (Hrsg.): *Marktforschung: Methoden, Anwendungen, Praxisbeispiele*, Gabler-Lehrbuch, Wiesbaden, S. 59–77.
- Schichl, H., R. Steinbauer* (2012): *Einführung in das mathematische Arbeiten*, 2. Aufl., Berlin u.a.
- Schneiderhan, V.* (2012): *Daseinsvorsorge und Vergaberecht: Darstellung eines Spannungsverhältnisses unter besonderer Berücksichtigung des Öffentlichen Personennahverkehrs*, Berlin.
- Schnell, R.* (1997): *Nonresponse in Bevölkerungsumfragen: Ausmaß, Entwicklung und Ursachen*, Opladen.
- Schnell, R., P.B. Hill, E. Esser* (2011): *Methoden der empirischen Sozialforschung*, 9. Aufl., München.
- Schnippe, C.* (1999): Relevanz von Qualitätskriterien: Der ÖPNV im Urteil der Fahrgäste, in: *Der Nahverkehr*, Jg. 17, Nr. 4, S. 52–56.
- Schnippe, C.* (2000): Psychologische Aspekte der Kundenorientierung: Die Kundenzufriedenheit mit der Qualität von Dienstleistungsinteraktionen am Beispiel des ÖPNV, *Europäische Hochschulschriften. Reihe 5, Volks- und Betriebswirtschaft*, Nr. 2663, Frankfurt am Main u.a.
- Scholl, A.* (2009): *Die Befragung*, 2. Aufl., Konstanz.
- Sedlmeier, P., F. Renkewitz* (2008): *Forschungsmethoden und Statistik in der Psychologie*, München.

- Severin, V., J.J. Louviere, A. Finn* (2001): The stability of retail shopping choices over time and across countries, in: *Journal of Retailing*, Vol. 77, No. 2, S. 185–202. doi:10.1016/S0022-4359(01)00043-4.
- Solaro, N.* (2012): Multidimensional Scaling, in: *Kenett, R.S., Salini, S. (Hrsg.): Modern analysis of customer surveys: With applications using R*, Chichester/UK, S. 357–390.
- Spencer, D.* (2009): Card sorting: Designing usable categories, Brooklyn/NY.
- Stevens, S.S.* (1946): On the Theory of Scales of Measurement, in: *Science*, Vol. 103, No. 2684, S. 677–680. doi:10.1126/science.103.2684.677.
- Stiefelbusch, M., M. Fliegel* (2006): Fahrgastbeiräte als Ratgeber für Verkehrsunternehmen: Konzepte und Erfahrungen in Deutschland und Österreich, in: *Der Nahverkehr*, Jg. 24, Nr. 7-8, S. 48–53.
- Stopher, P.R., Lee-Gosselin, M. (Hrsg.)* (1997): *Understanding travel behaviour in an era of change*, Oxford/UK.
- Sucharov, L.J. (Hrsg.)* (2002): *Urban transport VIII: Urban transport and the environment in the 21st century*, Southampton/UK u.a.
- Swait, J., A. Bernardino* (2000): Distinguishing taste variation from error structure in discrete choice data, in: *Transportation Research Part B: Methodological*, Vol. 34, S. 1–15. doi:10.1016/S0191-2615(99)00009-0.
- Swait, J., J.J. Louviere* (1993): The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models, in: *Journal of Marketing Research*, Vol. 30, S. 305–314, <http://www.jstor.org/stable/3172883> [Zuletzt geprüft am: 16.04.2013].
- Takane, Y., F. Young, J. de Leeuw* (1977): Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features, in: *Psychometrika*, Vol. 42, No. 1, S. 7–67. doi:10.1007/BF02293745.
- Theil, H.* (1971): *Principles of econometrics*, New York/NY u.a.
- Theobald, C., A. Kafka* (2003): ÖPNV-Finanzierung nach dem Urteil: EuGH-Beschluss stellt neue Anforderungen an Aufgabenträger und Verkehrsunternehmen, in: *Der Nahverkehr*, Jg. 21, Nr. 9, S. 11–13.
- Timmermans, H.J.P. et al.* (1992): Residential choice behaviour of dual earner households: a decompositional joint choice model, in: *Environment and Planning A*, Vol. 24, No. 4, S. 517–533. doi:10.1068/a240517.
- Timmermans, H.J.P., E.J.E. Molin* (2009): Dealing with Increased Complexity in Conjoint Experiments: Background and Overview of Alternate Approaches, in: *Transport Reviews*, Vol. 29, No. 5, S. 557–567. doi:10.1080/01441640902827532.
- TNS Infratest* (2012a): Das ÖPNV-Kundenbarometer 2012: Verkehrsverbünde und Verkehrsunternehmen im Vergleich - Die Spitzenreiter, München, http://www.tns-infratest.com/branchen_und_maerkte/pdf/verkehrsforschung/OPNV_Kundenbarometer_Ergebnisse.pdf [Zuletzt geprüft am: 16.04.2013].

- TNS Infratest* (2012b): ÖPNV-Kundenbarometer 2013: Der Schlüssel zu zufriedeneren Kunden und höheren Erträgen - Handlungsorientierte Analysen und Empfehlungen für höhere Kundenzufriedenheit und mehr Umsatz im ÖPNV, München, http://www.tns-infratest.com/branchen_und_maerkte/pdf/verkehrsforschung/TNS-Infratest_OPNV-Kundenbarometer-2013.pdf [Zuletzt geprüft am: 16.04.2013].
- Töpfer, A.* (Hrsg.) (1999): Kundenzufriedenheit messen und steigern, 2. Aufl., Neuwied.
- Train* (2001): A Comparison of Hierarchical Bayes and Maximum Simulated Likelihood for Mixed Logit, Berkeley/CA, <http://elsa.berkeley.edu/~train/compare.pdf> [Zuletzt geprüft am: 30.10.2012].
- Train, K.* (2009): Discrete choice methods with simulation, 2. Aufl., Cambridge/UK u.a.
- Transportation Research Board* (1999): A handbook for measuring customer satisfaction and service quality, Report 47, Washington, D.C., www.trb.org/publications/tcrp/tcrp_rpt_47-a.pdf [Zuletzt geprüft am: 16.04.2013].
- Ülengin, B.* (1998): Using hierarchical information integration to examine customer preferences in banking, in: *International Journal of Bank Marketing*, Vol. 16, No. 5, S. 202–210. doi:10.1108/02652329810228217.
- Ülengin, B., Ülengin Füsün, Ü. Gücedinç* (2002): Living Environment Preferences of the Inhabitants of Istanbul: A Modified Hierarchical Information Integration Model, in: *Social Indicators Research*, Vol. 57, S. 13–41. doi:10.1023/A:1013859201372.
- Ülengin, B., F. Ülengin, Ü. Güvenç* (2001): A multidimensional approach to urban quality of life: The case of Istanbul, in: *European Journal of Operational Research*, Vol. 130, No. 2, S. 361–374. doi:10.1016/S0377-2217(00)00047-3.
- Umweltbundesamt Dessau* (2005): Qualitätsziele und Indikatoren für eine nachhaltige Mobilität: Anwenderleitfaden, Dessau, <http://www.umweltdaten.de/publikationen/fpdf-l/2966.pdf> [Zuletzt geprüft am: 16.04.2013].
- van de Vyvere, Y., H. Oppewal, H.J.P. Timmermans* (1998): The Validity of Hierarchical Information Integration Choice Experiments to Model Residential Preference and Choice, in: *Geographical Analysis*, Vol. 30, No. 3, S. 254–272. doi:10.1111/j.1538-4632.1998.tb00399.x.
- van Helvoort-Postulart, D. et al.* (2009a): Discrete choice experiments for complex health-care decisions: does hierarchical information integration offer a solution?, in: *Health Economics*, Vol. 18, No. 8, S. 903–920. doi:10.1002/hec.1411.
- van Helvoort-Postulart, D. et al.* (2009b): Investigating the complementary value of discrete choice experiments for the evaluation of barriers and facilitators in implementation research: a questionnaire survey, in: *Implementation Science*, Vol. 4, No. 1, S. 10. doi:10.1186/1748-5908-4-10.

- Verband Deutscher Verkehrsunternehmen* (2012): VDV-Statistik 2011, Köln, http://mitglieder.vdv.de/module/layout_upload/st2011_online.pdf [Zuletzt geprüft am: 02.04.2013].
- Voeth, M.* (2000): Nutzenmessung in der Kaufverhaltensforschung: Die hierarchische individualisierte Limit-Conjoint-Analyse (HILCA), Wiesbaden.
- Vrtic, M., P. Fröhlich* (2006): Was beeinflusst die Verkehrsmittelwahl?: Erstellung eines ÖPNV-Verkehrsmodells für den Kanton Zürich, in: *Der Nahverkehr*, Jg. 24, Nr. 6, S. 52–57.
- Weis, H.C., P. Steinmetz* (2005): *Marktforschung*, 6. Aufl., Ludwigshafen am Rhein.
- Wilcoxon, F.* (1945): Individual Comparisons by Ranking Methods, in: *Biometrics Bulletin*, Vol. 1, No. 6, S. 80–83, <http://links.jstor.org/sici?sici=0099-4987%28194512%291%3A6%3C80%3AICBRM%3E2.0.CO%3B2-P> [Zuletzt geprüft am: 17.04.2013].
- Wolf, C., Best, H.* (Hrsg.) (2010): *Handbuch der sozialwissenschaftlichen Datenanalyse*, Wiesbaden.
- World Conference on Transport Research Society* (Hrsg.) (1992): *Traffic, Demand and Network Modelling*, Proceedings of the 6th World Conference on Transport Research, No. 2, Lyon.
- Wührer, G.A.* (2000): Mehrdimensionale Skalierung, in: Herrmann, A., Homburg, C. (Hrsg.): *Marktforschung: Methoden, Anwendungen, Praxisbeispiele*, Gabler-Lehrbuch, Wiesbaden, S. 439–469.
- Young, F.W., Y. Takane, R. Lewyckyj* (1978): Three notes on ALSCAL, in: *Psychometrika*, Vol. 43, No. 3, S. 433–435. doi:10.1007/BF02293652.
- Zarembka, P.* (Hrsg.) (1974): *Frontiers in econometrics*, New York/NY.

Quellenverzeichnis

- Allgemeines Eisenbahngesetz (AEG)* vom 27.12.1993 (BGBl. I S. 2378, 2396; 1994 I S. 2439), zuletzt geändert durch Artikel 2 des Gesetzes vom 14.12.2012 (BGBl. I S. 2598, 2606).
- DIN EN 13816*, Transport – Logistik und Dienstleistungen – Öffentlicher Personenverkehr – Definition, Festlegung von Leistungszielen und Messung der Servicequalität, hrsg. vom Deutschen Institut für Normung e.V., Juli 2002.
- Gesetz über den öffentlichen Personennahverkehr in Nordrhein-Westfalen (ÖPNVG NRW)* vom 7.3.1995 (GV. NW. 1995 S. 196), zuletzt geändert durch Gesetz vom 4.12.2012 (GV. NRW. S. 638).
- Gesetz über Finanzhilfen des Bundes zur Verbesserung der Verkehrsverhältnisse der Gemeinden (Gemeindeverkehrsfinanzierungsgesetz - GVFG)* vom 28.1.1988 (BGBl. I S. 100), zuletzt geändert durch Artikel 3 des Gesetzes vom 5.4.2011 (BGBl. I S. 554).
- Gesetz zur Entflechtung von Gemeinschaftsaufgaben und Finanzhilfen (Entflechtungsgesetz – EntflechtG)* vom 5.9.2006 (BGBl. I S. 2098, 2102).
- Gesetz zur Neuordnung des Eisenbahnwesens (Eisenbahnneuordnungsgesetz - ENeuOG)* vom 27.12.1993 (BGBl. I S. 2378; 1994 I S. 2439), zuletzt geändert durch Artikel 302 der Verordnung vom 31.10.2006 (BGBl. I S. 2407).
- Gesetz zur Regionalisierung des öffentlichen Personennahverkehrs (Regionalisierungsgesetz – RegG)* vom 27.12.1993 (BGBl. I S. 2378, 2395), zuletzt geändert durch Artikel 4 des Gesetzes vom 14.12.2012 (BGBl. I S. 2598).
- IBM Corporation* (2010): IBM SPSS Statistics for Windows. Version 19, Armonk/NY.
- IBM Corporation* (2011): IBM SPSS Statistics for Windows. Version 20, Armonk/NY.
- StataCorp.* (2011): Stata Statistical Software. Release 12, College Station/TX.
- Verordnung (EG) Nr. 1370/2007* des Europäischen Parlamentes und des Rates vom 23. Oktober 2007 über öffentliche Personenverkehrsdienste auf Schiene und Straße und zur Aufhebung der Verordnungen (EWG) Nr. 1191/69 und (EWG) Nr. 1107/70 des Rates (ABl. Nr. L 315/1 vom 3.12.2007).

Einfluss der Qualität des ÖPNV auf die Verkehrsmittelwahl im Regionalverkehr

Cornelia Richter

In dieser Arbeit wird der Zusammenhang zwischen Qualität und Verkehrsmittelwahl untersucht. Dazu werden Discrete Choice Experimente mit über 2000 Nutzern und Nicht-Nutzern des regionalen ÖPNV durchgeführt. Durch die Verwendung des Verfahrens der integrierten hierarchischen Informationsintegration (HII-I) kann eine größere Anzahl von Qualitätsattributen in den Experimenten berücksichtigt werden, da hierbei ähnliche Attribute zu Konstrukten zusammengefasst werden. Diese Konstrukte werden im Rahmen der Arbeit empirisch hergeleitet.

Mit zahlreichen statistischen Tests wird der Einfluss der Qualität auf die Verkehrsmittelwahl, Unterschiede zwischen einzelnen Nutzergruppen sowie die hierarchische Struktur des eingesetzten Verfahrens überprüft.

Aufbauend auf den Ergebnissen dieser Arbeit lassen sich Veränderungen des Modal Splits infolge einer Änderung der Qualität simulieren und Empfehlungen für die Praxis ableiten.

ISBN 978-3-8405-0094-7

EUR 19,40

