

# METHODISCHE GRUNDLAGEN DES DFG-PROJEKTS "WOHNSTANDORTWAHL, RAUM UND VERKEHR IM KONTEXT VON LEBENSSTIL UND LEBENSLAGE"

Datenerhebung, Datenaufbereitung, Datenanalyse

von JOACHIM SCHEINER



Dr. Joachim Scheiner  
Technische Universität Dortmund  
Fakultät Raumplanung  
Fachgebiet Verkehrswesen und Verkehrsplanung  
D-44221 Dortmund  
Tel 0231/755-4822  
Fax 0231/755-2269  
[joachim.scheiner@tu-dortmund.de](mailto:joachim.scheiner@tu-dortmund.de)

## Inhalt

<b>1</b>	<b>EINLEITUNG.....</b>	<b>1</b>
<b>2</b>	<b>PROJEKTANSATZ.....</b>	<b>1</b>
<b>3</b>	<b>DATENERHEBUNG.....</b>	<b>3</b>
3.1	Haushaltsbefragung des Projekts StadtLeben.....	3
3.2	Raumstrukturelle Daten.....	4
3.2.1	Arbeitsplatzangebot.....	4
3.2.2	Versorgungs- und Freizeitgelegenheiten, ÖPNV.....	5
3.2.3	Schulen.....	6
<b>4</b>	<b>DATENAUFBEREITUNG.....</b>	<b>6</b>
4.1	Behandlung fehlender Werte.....	6
4.2	Lebenslage I.....	8
4.3	Lebenslage II.....	9
4.4	Lebensstil I.....	10
4.4.1	Überblick über die Vorgehensweise.....	10
4.4.2	Schritt 1: Faktorenanalytische Bildung von Lebensstildimensionen.....	11
4.4.3	Schritt 2: Bildung von jeweils an ein Modell inhaltlich angepassten Faktoren.....	13
4.5	Lebensstil II.....	16
4.6	Wohnstandortpräferenzen.....	16
4.7	Objektive Raumstruktur.....	17
4.7.1	Arbeitsplatzangebot.....	17
4.7.2	Schulen.....	20
4.7.3	Versorgungs- und Freizeitgelegenheiten.....	20
4.7.4	ÖPNV-Angebot.....	20
4.7.5	Lage der Gebiete.....	21
4.7.6	Bildung raumstruktureller Indikatoren.....	21
4.8	Wohnmobilität und Standortwahl.....	23
4.9	Verkehrsnachfrage.....	25
<b>5</b>	<b>METHODISCHE ANALYSEN: EVALUATION DER RAUMSTRUKTURELLEN INDIKATOREN.....</b>	<b>26</b>
5.1	Lohnt die personenbezogene Differenzierung der Wohnstandorte innerhalb der Gebiete?.....	26
5.2	Lohnt der Aufwand der Bildung eines Indikators 'Arbeitsplatzerreichbarkeit'?.....	28
<b>6</b>	<b>ANALYSEVERFAHREN.....</b>	<b>29</b>
6.1	Methodik linearer Strukturgleichungsmodelle.....	29
6.2	Verteilung der Variablen und Skalierung.....	31
6.3	Angewandtes Schätzverfahren und Validierung der Ergebnisse.....	31
6.4	Ausblick.....	32
<b>7</b>	<b>LITERATUR.....</b>	<b>33</b>

## METHODISCHE GRUNDLAGEN DES DFG-PROJEKTS "WOHNSTANDORTWAHL, RAUM UND VERKEHR IM KONTEXT VON LEBENSSTIL UND LEBENSLAGE" - Datenerhebung, Datenaufbereitung, Datenanalyse

JOACHIM SCHEINER

**Zusammenfassung:** Das durch die DFG geförderte Forschungsprojekt "Wohnstandortwahl, Raum und Verkehr im Kontext von Lebensstil und Lebenslage" ist der empirischen Analyse der komplexen Zusammenhänge zwischen Lebenslage, Lebensstil, Wohnstandortpräferenzen, Wohnstandortwahl und Verkehrsnachfrage gewidmet. Die Analysen basieren insbesondere auf linearen Strukturgleichungsmodellen. Das Arbeitspapier präsentiert und diskutiert methodische Grundlagen der Datenerhebung, Datenaufbereitung und Datenanalyse.

**Summary:** The research project 'Choice of residential location, the built environment and transport in the context of lifestyle and life situation', which is funded by the German Research Foundation DFG, is devoted to the empirical analysis of the complex interrelations between life situations, lifestyles, location attitudes, residential location choice, and travel behaviour. The analyses are based on linear structural equation modelling. The paper presents and discusses methodological considerations on the data survey, data processing and data analysis.

### 1 Einleitung

Das durch die Deutsche Forschungsgemeinschaft DFG geförderte Forschungsprojekt "Wohnstandortwahl, Raum und Verkehr im Kontext von Lebensstil und Lebenslage" ist der empirischen Analyse der komplexen Zusammenhänge zwischen Lebenslage, Lebensstil, Wohnstandortpräferenzen, Wohnstandortwahl und Verkehrsnachfrage gewidmet. Die Analysen stützen sich auf eine Haushaltsbefragung in zehn Untersuchungsgebieten der Region Köln, die im Rahmen des Projekts StadtLeben unternommen wurde. Die Daten dieser Befragung werden vorwiegend mit Hilfe linearer Strukturgleichungsmodelle analysiert. Dieses Papier stellt methodische Grundlagen der Datenerhebung, Datenaufbereitung und Datenanalyse vor.

Im Folgenden wird zunächst der theoretische Projektansatz kurz vorgestellt. Danach wird ein kurzer Überblick über die erhobenen Daten gegeben. Anschließend werden die umfangreichen Schritte der Datenaufbereitung dargestellt. Dies umfasst auch Variablen, die in dem DFG-Projekt nicht zum Einsatz kommen, aber evtl. für zukünftige Analysen von Bedeutung sind. In Kapitel 5 werden einige Überlegungen zu den differenzierten raumstrukturellen Indikatoren dargestellt, da speziell deren Erhebung und Aufbereitung umfangreiche Arbeitsschritte erforderte, so dass sich die Frage stellt, ob dies erforderlich ist. Abschließend folgen einige Überlegungen zum Analyseverfahren einschließlich der Validierung der Ergebnisse.

### 2 Projektansatz

Im Folgenden wird der theoretische Ansatz des Projekts vorgestellt. Dieser beruht auf fünf Kernbausteinen: Lebenslage, Lebensstil, Standortpräferenzen, Standortwahl/Raumstruktur und Verkehrshandeln. Der Ansatz basiert auf Überlegungen, die sich im Kontext der BMBF-Projekte StadtLeben und Inter-

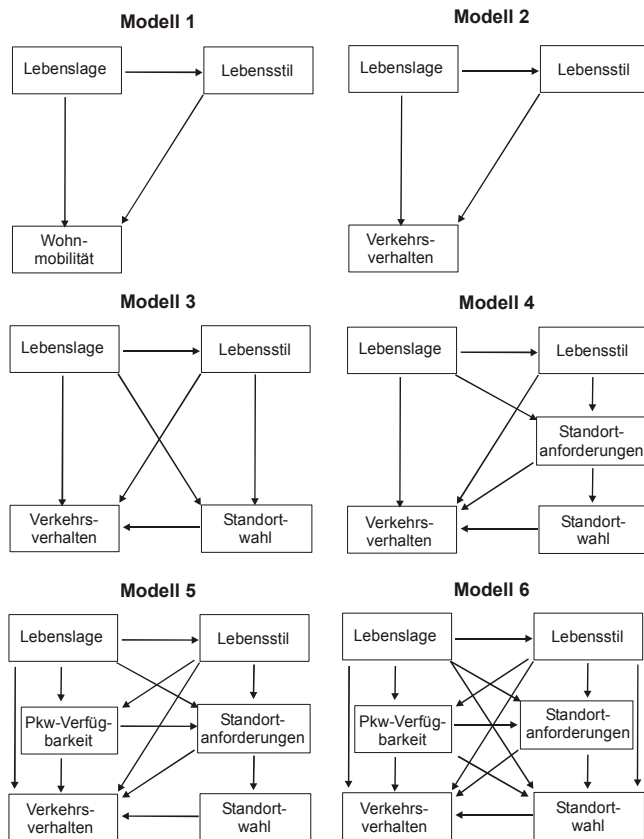
bil Region Dresden entwickelt haben<sup>1</sup>. Diese werden hier nur kurz zusammengefasst. Ausführliche Darstellungen finden sich in Scheiner/Kasper (2005), Hammer/Scheiner (2006), Scheiner (2006b) und Scheiner/Holz-Rau (2007).

Zunächst postuliert der vorgestellte Ansatz einen Einfluss des Lebensstils auf räumliche Mobilität (Wohnmobilität, Verkehrsnachfrage). Gleichzeitig wird in Übereinstimmung mit vielen soziologischen Studien behauptet, dass Lebensstile ihrerseits von sozialen und demografischen Strukturen, kurz von der Lebenslage, abhängig sind (Schneider/Spellerberg 1999). Im Gegensatz zu manchen anderen Lebensstilansätzen werden hier Lebensstil und Lebenslage strikt getrennt. Während Lebensstile stets ein Moment der Handlungsfreiheit und Freiwilligkeit ("Stilisierung") besitzen und "zwischen sozialer Lage und individuellem Handeln, objektiven Lebensbedingungen und kulturellem Leben" (Spellerberg 1996:53) vermitteln, stellen Lebenslagen in erster Linie objektive, strukturelle Ungleichheiten dar. Diese lassen sich mittels sozioökonomischer und demografischer Merkmale beschreiben, etwa Einkommen, Bildung, Geschlecht, Alter, Nationalität oder Kohortenzugehörigkeit (Berger/Hradil 1990:10). Analytisch

---

<sup>1</sup> "StadtLeben – Integrierte Betrachtung von Lebensstilen, Wohnumlieus, Raum- und Zeitstrukturen für die zukunfts-fähige Gestaltung von Mobilität und Stadt". Laufzeit 2001-2005. Projektpartner: RWTH Aachen, Institut für Stadtbauwesen und Stadtverkehr (Koordination); FU Berlin, Institut für Geographische Wissenschaften, Abt. Stadtforschung; Ruhr-Universität Bochum, Arbeitseinheit Kognitions- und Umweltpsychologie; Technische Universität Dortmund, Fachgebiet Verkehrswesen und Verkehrsplanung (näheres unter <http://www.isb.rwth-aachen.de/stadtleben/>). – "Intermobil Region Dresden" (Laufzeit 1999-2004) war ein großes Verbundprojekt mit einer Vielzahl von Partnern. Der Autor war als Beschäftigter des Büros für Integrierte Planung am Teilprojekt "Raum- und verhaltensstrukturelle Voraussetzungen der nachhaltigen Mobilitätssicherung" beteiligt, das von der Technischen Universität Hamburg-Harburg, Arbeitsbereich Verkehrssysteme und Logistik, geleitet wurde (Projektinformationen unter <http://www.vsl.tu-harburg.de/100>). Beide Projekte wurden durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert.

sollten Lebenslage und Lebensstil als 'objektive' und 'subjektive' Ungleichheitsmerkmale nicht vermischt werden, damit entsprechende Zusammenhänge explizit geprüft werden können (Wieland 2004:176f). Nur dann ist auch der Mehrwert von Lebensstilen gegenüber der Lebenslage zur Erklärung räumlicher Mobilität prüfbar. Daraus ergeben sich die Strukturen der Modelle 1 und 2, die in Abbildung 1 dargestellt sind.



**Abbildung 1: Modellstrukturen zur Abbildung von Lebenslage, Lebensstil, Standortpräferenzen, Standortwahl und Verkehrshandeln**

Quelle: eigener Entwurf

Des Weiteren wird behauptet, dass zwischen Wohnmobilität und Verkehrsnachfrage wechselseitige Abhängigkeitsstrukturen bestehen. So lässt sich zeigen, dass die höhere Motorisierung in Suburbia gegenüber den Kernstädten einerseits aus Selbstselektionseffekten der Bevölkerung herrührt: Randwanderer sind bereits vor der Abwanderung aus der Stadt höher motorisiert als sesshafte Städter. Andererseits erhöht sich die Motorisierung der entsprechenden Haushalte nach der Randwanderung weiter. Es erfolgt also eine Anpassung an die pkw-orientierten Raumstrukturen Suburbias (Scheiner 2005a).

Rückwirkungen sollen allerdings in der Modellierung ausgeklammert bleiben, weil lediglich Querschnittdaten zur Analyse vorliegen und Wechselwirkungen nur aus der zeitlichen Abfolge von Ereignissen erschlossen werden können. Wir nehmen also einsei-

tige Ursache-Wirkungs-Beziehungen an. Dabei kommt den Auswirkungen der Wohnmobilität als langfristiger Entscheidung auf das Verkehrshandeln Priorität gegenüber der umgekehrten Wirkungsbeziehung zu. Der wichtigste Aspekt der Wohnmobilität ist hier die Standortwahl. Daraus ergibt sich die in Modell 3 dargestellte Modellstruktur.

Weiterhin können wir davon ausgehen, dass Standortentscheidungen von bestimmten Wohn- und Standortpräferenzen bestimmt werden. Darunter ist die subjektive Wichtigkeit von Merkmalen der Wohnung und des Standorts für eine Person zu verstehen. Im Projekt StadtLeben wurde hier von Standortanforderungen gesprochen (Scheiner 2006a). Mit Blick darauf, dass es sich hierbei nicht unbedingt auch um objektive Anforderungen im Sinne faktisch benötigter Eigenschaften der Wohnung oder des Standorts handelt, wird hier von (subjektiven) Präferenzen gesprochen<sup>2</sup>. Merkmale der Wohnung interessieren hier nur indirekt aufgrund der ungleichen räumlichen Verteilung verschiedener Wohnungsgrößen, -schnitte, -ausstattungen, Haustypen und Eigentumsverhältnisse. Sie werden im Folgenden nicht weiter betrachtet. Deshalb wird hier zusammenfassend von "Standortpräferenzen" gesprochen.

Diese Standortpräferenzen sind teilweise durch die Lebenslage bestimmt, etwa durch das Vorhandensein von Kindern im Haushalt, können aber auch durch individuelle Werte und Präferenzen bestimmt werden, die durch den Lebensstil beschrieben werden. Wir können also Standortpräferenzen als weitere intervenierende Variablen zwischen Lebenslage und Lebensstil einerseits und Standortwahl andererseits berücksichtigen (Modell 4).

In einem weiteren Schritt wird die Pkw-Verfügbarkeit integriert (Modell 5), weil sie – ähnlich wie die Standortwahl – eine wichtige Vorentscheidung für das Verkehrshandeln darstellt. Die Pkw-Verfügbarkeit ist von den materiellen Ressourcen einer Person bzw. eines Haushalts abhängig und kann insofern als intervenierende Größe zwischen Lebenslage und Verkehrshandeln gelten. Potenziell ist sie auch vom Lebensstil abhängig. Die durch den Pkw erzielte höhere zeit-räumliche Erreichbarkeit führt dazu, dass die Pkw-Verfügbarkeit auch die Standortpräferenzen einer Person mitbestimmt (z.B. geringere Bedeutung von Nähe zu Versorgungseinrichtungen) und vermittelt darüber auch indirekte Effekte auf das Verkehrshandeln ausüben kann.

In einem letzten Schritt schließlich (Modell 6) können auch direkte Effekte der Lebenslage, des Lebensstils und der Pkw-Verfügbarkeit auf die Standortwahl

<sup>2</sup> In der thematisch verwandten englischsprachigen Wohnstandort- und Verkehrsforschung werden die Begriffe "attitude" und "preference" gleichermaßen verwendet. Einstellung (attitude) ist allerdings etwas unscharf, weil Einstellungen auch Zufriedenheiten umfassen.

untersucht werden, da nicht davon ausgegangen werden kann, dass die Standortwahl vollständig von den Standortpräferenzen einer Person bestimmt wird. Erstens können Standortpräferenzen – wie auch die anderen Grundbausteine der Modelle – aus Gründen der Überschaubarkeit hier stets nur partiell berücksichtigt werden. Zweitens sind Standortentscheidungen Haushaltsentscheidungen, die nicht unbedingt den Standortpräferenzen jeder Person im Haushalt entsprechen. Haushaltsstrukturen aber kommen in der Lebenslage zum Ausdruck. Drittens werden Standortentscheidungen in hohem Maß durch die vorherrschenden Marktverhältnisse (Verfügbarkeit und Preis von Wohnraum) bestimmt.

Innerhalb der beschriebenen Grundstrukturen sind jeweils eine Vielzahl von Fragestellungen analysierbar, die wiederum jeweils unterschiedlich operationalisiert werden können. So wird das Verkehrshandeln im Rahmen des Projekts unter den Aspekten Aktivitätshäufigkeit, Wegelängen (Distanzen) und Verkehrsmittelnutzung untersucht. Diese Indikatoren sind auch in Kombination miteinander in die Modelle integrierbar. Beispielsweise hängen Wegelänge und Verkehrsmittelnutzung eng miteinander zusammen. So kommen nicht-motorisierte Verkehrsmittel nur für relativ kurze Wege in Frage. Die Verkehrsmittelnutzung kann also in Abhängigkeit von der Wegelänge modelliert werden. Auch das umgekehrte Vorgehen ist plausibel, bei der die Wegelänge in Abhängigkeit von der Verkehrsmittelnutzung modelliert wird: Wer bereits im Auto sitzt, für den spielt die Entfernung – in gewissem Rahmen – keine so große Rolle mehr.

### 3 Datenerhebung

#### 3.1 Haushaltsbefragung des Projekts StadtLeben

Die im Rahmen des Projekts analysierten Daten wurden im Rahmen einer standardisierten Haushaltsbefragung des Projekts StadtLeben erhoben (Beckmann/Hesse/Holz-Rau/Hunecke 2006). Die Befragung fand in den Jahren 2002 und 2003 in zehn Untersuchungsgebieten der Region Köln statt. Dabei wurden  $n=2.691$  Bewohner in ausführlichen persönlichen Interviews zu ihrem Verkehrshandeln, zur Wohnmobilität, zu Lebenslage und Lebensstilen, zu Wohn- und Standortpräferenzen und Standortzufriedenheit befragt. Die Ausschöpfungsquoten liegen je nach Gebiet zwischen 23 % und 37 % der bereinigten Bruttostichprobe (insgesamt 27%). Angesichts des hohen Aufwands für die Befragten (ca. 1 Stunde Befragungsdauer) erscheint dies angemessen.

Die Untersuchungsgebiete stehen für fünf Gebietstypen, die durch jeweils zwei Gebiete repräsentiert werden: hochverdichtete Innenstadtgebiete der Gründerzeit (Ehrenfeld, Nippes), durchgrünte Zeilenbaugebiete der Nachkriegszeit in Innenstadtrandlage

(Stammheim, Longerich), periphere Wohngebiete am Rand der Kernstadt Köln (Esch, Zündorf), Mittelzentren des suburbanen Umlands (Kerpen Stadt, Overath Stadt) und periphere Wohngebiete des suburbanen Umlands (Kerpen-Sindorf, Overath-Heiligenhaus). Da auch die beiden Gebiete eines Typs sich jeweils deutlich unterscheiden, umfassen die Gebiete eine recht große Spannweite im Hinblick auf räumliche Lage, Ausstattung mit Verkehrsinfrastruktur, zentralörtliche Einrichtungen und soziale Strukturen. Gleichwohl ist zu betonen, dass nicht versucht wurde, soziale oder räumliche "Extremgebiete" auszuwählen. Es sind keine ausgesprochenen Oberschichtgebiete und nur ein deutlich als Unterschichtquartier erkennbares Gebiet (Stammheim) vertreten.

Drei Gebiete (Stammheim, Ehrenfeld und Esch) werden in den Analysen nicht berücksichtigt, weil dort aus Gründen des Projektablaufs die Standortpräferenzen der Befragten nicht erfragt werden konnten. Damit basieren die Analysen lediglich auf sieben Untersuchungsgebieten. Die resultierenden Netto-Stichproben liegen je nach Modell bei rund  $n=2.000$ .

In räumlicher Hinsicht repräsentiert Heiligenhaus das am stärksten als peripher zu kennzeichnende Quartier; dabei handelt es sich um ein reines Wohngebiet ohne nennenswerte Ausstattung mit Versorgungseinrichtungen und sehr ungünstiger ÖPNV-Anbindung. Aber auch Heiligenhaus liegt im hochverdichteten Umland der Großstadt Köln, ist also – im Sinne der Gesamtspannweite der Raumstruktur Deutschlands – nicht ausgesprochen peripher.

Die Region Köln ist ein polyzentraler Verdichtungsraum mit dem klar dominierenden Zentrum Köln. Die Bevölkerungsentwicklung ist leicht positiv und der Wohnungsmarkt wird eher von der Anbieterseite dominiert. Die Chancen unterschiedlicher Lebensstil- und Lebenslagegruppen, einen ihren Standortpräferenzen entsprechenden Wohnstandort realisieren zu können, sind also eher eingeschränkt. Dies ist eine für die Interpretation der Ergebnisse wichtige Rahmenbedingung.

Meines Wissens sind die verwendeten Daten einzigartig darin, alle für die beschriebenen Modellstrukturen notwendigen Daten auf der Individualebene verknüpfbar zu machen. Dennoch besitzen auch diese Daten Mängel, deren wichtigste an dieser Stelle erwähnt werden sollen.

1. Aufgrund des Querschnittscharakters der Daten (mit einigen retrospektiven Elementen) kann sich die Ableitung von Kausalitäten der ermittelten Zusammenhänge nicht auf zeitliche Abläufe, sondern nur auf Plausibilitätsüberlegungen bzw. theoretische Überlegungen stützen. Der Fragebogen umfasst allerdings einige retrospektive Fragen, die zur Konstruktion von Quasi-Panel-Analysen genutzt werden können. Dies betrifft insbesondere Haushaltsstruktur, Pkw-Verfügbar-

keit und Wohnstandort vor dem letzten Umzug sowie Veränderungen der Verkehrsmittelnutzung seit dem Umzug.

2. Alle Daten wurden bezogen auf den Zeitpunkt der Befragung erhoben, unterliegen aber dennoch unterschiedlichen Fristigkeiten und zeitlichen Abfolgen. Der letzte Umzug liegt beispielsweise in der Vergangenheit. Die Standortentscheidung wurde also in einer bestimmten Lebenslage getroffen, die nicht der Lebenslage entspricht, die für das Verkehrshandeln die Grundlage bildet. Inwieweit die Lebensstile und individuellen Standortpräferenzen sich seit dem Umzug geändert haben, ist vollkommen unbekannt.
3. Den Modellen liegen sowohl individuelle Handlungen (Verkehrshandeln) als auch Handlungen von Haushalten (Umzug) zugrunde. Die daraus resultierenden potenziellen Abhängigkeiten sind aufgrund der personenbezogenen Befragung nicht in die Modelle integrierbar. So ist es etwa vorstellbar, dass die Standortpräferenzen im Modell keinen Einfluss auf die tatsächliche Standortentscheidung besitzen, weil die Entscheidung auf der Grundlage der Standortpräferenzen des Partners getroffen wurde. Die ermittelten empirischen Zusammenhänge – hier zwischen Standortpräferenzen und Standortentscheidung – dürften insoweit eher konservative Schätzungen darstellen, die in der Realität stärker ausfallen dürften.

## 3.2 Raumstrukturelle Daten

### 3.2.1 Arbeitsplatzangebot

Das Arbeitsplatzangebot in einem Gebiet lässt sich im Gegensatz etwa zum Einzelhandelsangebot durch Vor-Ort-Erhebungen kaum ermitteln, weil man einem Gebäude die Zahl der darin befindlichen Arbeitsplätze nicht ansieht und eine Betriebsbefragung aus forschungsökonomischen und anderen Gründen kaum praktikabel ist. Zudem ist für die Erreichbarkeit von Arbeitsmärkten nicht nur das Arbeitsplatzangebot in einem kleineren Radius relevant, sondern auch das Angebot im weiteren Umfeld. Demzufolge müsste eine angemessene Erhebung des Arbeitsplatzangebots ein wesentlich größeres Gebiet umfassen<sup>3</sup>.

Deshalb wird zur Ermittlung der Arbeitsplatzreichtbarkeit auf vorhandene Daten zurückgegriffen. Aus

---

<sup>3</sup> So werden im Berufsverkehr aufgrund der Spezialisierung in der Arbeitswelt im Mittel deutlich höhere Distanzen je Weg zurückgelegt als im Einkaufs- und Freizeitverkehr. Aus Verkehr in Zahlen ergibt sich für 2002 ein Wert von 14,2 km je Weg im Berufsverkehr, 12,3 km im Freizeitverkehr und 5,8 km im Einkaufsverkehr. Im Freizeitverkehr ergeben sich die recht großen Mittelwerte vor allem aus wenigen sehr langen Wegen, etwa für Verwandtenbesuche.

amtlichen Daten steht allerdings das Arbeitsplatzangebot lediglich als Zahl der sozialversicherungspflichtigen Beschäftigten (gezählt am Arbeitsort) auf der Gemeindeebene zur Verfügung. Dies ist für eine regionale Studie nicht ausreichend, zumal sechs der zehn Untersuchungsgebiete in Köln liegen und somit das erreichbare Arbeitsplatzangebot auf der Gemeindeebene sich für diese sechs Gebiete nur geringfügig unterscheidet. Eine Anfrage bei der Bundesanstalt für Arbeit nach einer kleinräumlichen Disaggregation für die Stadt Köln (etwa auf Ebene der Postleitzahlbezirke) wurde abschlägig beantwortet. Die Stadt Köln selbst verfügt ebenfalls nicht über derartige Zahlen.

Kleinräumlich disaggregierte Werte für Köln wurden schließlich von der Firma PTV zur Verfügung gestellt<sup>4</sup>. Dabei handelt es sich um Werte auf der Ebene der 721 Kölner Verkehrszellen, die ursprünglich auf die Volkszählung 1987 zurückgehen, jedoch durch PTV in Kooperation mit dem Amt für Stadtentwicklung und Statistik der Stadt Köln auf etwa das Jahr 2000 fortgeschrieben wurden. Die Fortschreibung erfolgte räumlich differenziert. Die Basis dafür bildeten sowohl amtliche Daten als auch eigene Ortskenntnisse, z.B. über die räumliche Verteilung der Arbeitsplätze von Arbeitgebern mit mehreren Betrieben, für die nur am Hauptsitz amtliche Daten vorlagen. Diese Daten wurden anhand des amtlichen Summenwertes für die Stadt Köln in 2003 leicht nach oben angepasst.

Für die Städte Kerpen und Overath lagen keine entsprechenden Daten vor. Die räumliche Disaggregation wurde deshalb auf andere Weise vorgenommen. Anhand der privatwirtschaftlich erstellten Datenbank "Markus" (Bureau van Dijk 2004), in der deutsche Unternehmen mit Anzahl der Beschäftigten und Adresse gelistet sind, wurde eine räumliche Verteilung der Beschäftigten in Kerpen und Overath auf die Ortsteile vorgenommen und in der Summe an den amtlichen Werten der sozialversicherungspflichtig Beschäftigten (2003) für die Gesamtgemeinden justiert. Die genannte Datenbank zählt leider alle Beschäftigten eines Unternehmens am Standort des Hauptsitzes, so dass die Beschäftigten von Mehrbetriebsunternehmen teilweise am falschen Standort gezählt werden. Deshalb wurde die ermittelte räumliche Verteilung mit Vertretern der Städte Kerpen und Overath auf ihre Plausibilität hin diskutiert<sup>5</sup>. In beiden Städten wurde die ermittelte räumliche Verteilung für das Jahr 2003 als plausibel eingeschätzt. In Kerpen war dies insbesondere für den herausragenden Wert im Ortsteil Sindorf wichtig, wo sich rund 60% der Arbeitsplätze Kerpens befinden.

---

<sup>4</sup> Mein Dank geht an dieser Stelle an Dipl.-Ing. Gunther Kesenheimer, PTV, Karlsruhe.

<sup>5</sup> Mein Dank gilt an dieser Stelle den Herren Bernd Sassenhof, Beigeordneter der Stadt Overath, und Andreas Comacchio, Wirtschaftsförderung der Stadt Kerpen.

Ergänzt wurden die kleinräumlichen Daten für die Untersuchungsgemeinden durch die amtlichen Daten der sozialversicherungspflichtig Beschäftigten für das gesamte Land Nordrhein-Westfalen auf der Gemeindeebene.

### 3.2.2 Versorgungs- und Freizeitgelegenheiten, ÖPNV

Die Erreichbarkeit von Versorgungs- und Freizeitgelegenheiten sowie ÖPNV-Angeboten gründet sich vor allem auf die Ausstattung der Wohnumgebung mit entsprechenden Gelegenheiten.

Im Rahmen des Projekts StadtLeben wurde eine umfangreiche Bestandsaufnahme der Flächennutzung in den Untersuchungsgebieten vorgenommen. Diese umfasst insbesondere einen Behavior Setting Survey der Ruhr-Universität Bochum und Kartierungen der RWTH Aachen. Aus diesen Erhebungen liegen geographische Koordinaten von Versorgungs- und Freizeitangeboten sowie Angaben über die Art des Angebots vor, außerdem Koordinaten der ÖPNV-Haltepunkte sowie eine Variable, die das ÖPNV-Angebot in den Gebieten beschreibt. Die Erhebung von Freizeitangeboten umfasst neben baulichen Einrichtungen auch informelle Treffpunkte und Aktivitätsorte, die im Rahmen des erwähnten Behavior Setting Survey erhoben wurden.

Diese Daten werden ergänzt durch eigene Vor-Ort-Erhebungen. Diese wurden im Frühjahr 2006 durch cand.-ing. Julia Hampe und cand.-ing. Janina Kranicz durchgeführt. Sie dienen dazu, den Radius der Erhebungsgebiete auszudehnen. Im Projekt StadtLeben erwies es sich als nützlich, mit personenbezogenen Variablen der Erreichbarkeit von Versorgungs- und Freizeiteinrichtungen zu arbeiten (Huncke/Schweer 2006). Da aber aus forschungsökonomischen Gründen Behavior Setting Survey und Kartierung nicht überall in hinreichendem Maß über die Grenzen der Befragungsgebiete hinaus ausgedehnt werden konnten, war die Validität der entsprechenden Indikatoren für befragte Personen in den Randbereichen der Befragungsgebiete nicht gesichert. Durch die ergänzenden Erhebungen über die Grenzen der Befragungsgebiete hinaus wird sichergestellt, dass für jede befragte Person Versorgungs- und Freizeiteinrichtungen im Radius von 650 m Luftlinie um den Wohnstandort auch außerhalb der Befragungsgebiete vorliegen. Die erhobenen Einrichtungen umfassen Einzelhandel, Dienstleistungen, Freizeiteinrichtungen, soziale Einrichtungen und ÖPNV-Haltepunkte. Informelle Treffpunkte wurden ebenfalls erhoben, allerdings mit geringerem Aufwand als im Projekt StadtLeben. Konkret fand nur eine einmalige Begehung statt, so dass temporäre Nutzungen mit hoher Wahrscheinlichkeit nicht vollständig erfasst wurden.

Die Vor-Ort-Erhebung erfolgte differenziert nach Art und Größe sowie mit Adresse. Die Adresse ist notwendig zur Ermittlung der geographischen Koordinaten

und zur anschließenden Bildung personenbezogener Indikatoren (z.B. "wie viele Einzelhandelseinrichtungen befinden sich im Umkreis von 200 m um den Wohnstandort eines Befragten?"). Der Radius von 650 m Luftlinie um die Befragungsgebiete gewährleistet, dass die personenbezogene Erreichbarkeit von Gelegenheiten im Radius von 650 m Luftlinie auch für befragte Personen, die am äußersten Rand eines Untersuchungsgebiets leben, vollständig berechnet werden kann. Dies entspricht in etwa der maximal akzeptierten Fußwegelänge bei Einkaufswegen ( $650 \text{ m} \cdot \text{Umfwegfaktor } 1,2 = 780 \text{ m}$ )<sup>6</sup>. Im Freizeitverkehr werden insgesamt deutlich längere Wege in Kauf genommen als im Einkaufsverkehr; das großräumige Freizeitangebot ist jedoch nicht adäquat über vorliegende Indikatoren abbildbar und durch eigene Erhebungen nicht zu bewältigen.

Der aus den verschiedenen Erhebungen resultierende Datensatz umfasst rund 3.800 Gelegenheiten, von denen etwa 2.900 zu verschiedenen Erreichbarkeitsindikatoren weiterverarbeitet wurden. Die nicht weiterverarbeiteten Einrichtungen stammen zumeist aus den Erhebungen des Projekts StadtLeben. Sie umfassen vorwiegend Handwerksbetriebe, unternehmensbezogene Dienstleistungen, Großhandel, Lagerhaltung, Betreuungseinrichtungen (Altenheime, Kindergärten) und leerstehende Gebäude, die für die Erstellung von Erreichbarkeitsindikatoren der erfragten Aktivitäten als nicht wesentlich angesehen wurden.

Für den Bereich Versorgung ist eine Kartierung vergleichsweise unproblematisch. In der Freizeit ist allerdings ein erheblicher Teil der Mobilität auf Gelegenheiten gerichtet, die nicht durch 'gewöhnliche' Kartierungen zu ermitteln sind. Dazu zählen neben informellen Treffpunkten insbesondere auch private Wohnungen. Nach den Daten von 'Mobilität in Deutschland' 2002 (eigene Analyse) stellen rund 30% der Wege in der Freizeit private Besuche dar. Dabei sind Freizeitwege mit Zwecken wie 'Essen', die teilweise ebenfalls in privaten Wohnungen stattfinden dürften, noch nicht mitgerechnet<sup>7</sup>. Auch die Eignung eines Quartiers für Spaziergänge ist nicht ohne Weiteres zu bestimmen. Grünflächen können zwar kartiert werden, unterscheiden sich aber stark in ihrer Qualität, und auch die bebaute Umgebung ist qualitativ in sehr unterschiedlichem Maß für Spazierwege geeignet. Diesen Überlegungen folgend, wurden folgende weitere Daten verwendet.

Die Verteilung von Potenzialen für private Besuche wurde, ähnlich der Vorgehensweise in einem Ver-

<sup>6</sup> Nach einer kleinräumlichen Studie von Holz-Rau (1991) werden Einkaufsverkehr bereits ab einer Entfernung von 600-700 m überwiegend mit dem Pkw zurückgelegt, sofern ein solcher verfügbar ist.

<sup>7</sup> Nach Zängler (2000, S. 229) sind etwa 40% der Wege in der Freizeit private Besuche.

kehrsberechnungsmodell, mit Hilfe eines Gravitationsansatzes auf der Grundlage der räumlichen Bevölkerungsverteilung ermittelt.

Gebiet	Beschreibung	Bewertung
Ehrenfeld	Hohe Dichte, laut, eng; Grüngürtel östlich Ehrenfeld zum Spazieren, Ausruhen, Picknick etc. geeignet. Melaten-Friedhof allenfalls zum Spazieren geeignet	1
Nippes	ähnlich Ehrenfeld, aber höhere "urbane Qualität"; im Nordpark soziale Probleme; mit Fahrrad sind Zoo und Rheinufer gut erreichbar	2
Stammheim	Rheinufer und Schlosspark bieten hohe Qualität für Aufenthalt, Spazieren und Verweilen; Gastronomie am Ufer vorhanden; Fußweg zur Innenstadt; soziale Probleme	6
Longerich	zusammenhängende Grünflächen vorhanden, die aber aufgrund sozialer Probleme als 'unbehaglich' gelten (kaputte Bänke, fehlendes Licht, 'Zigeunersiedlung'); besondere Attraktivität (z.B. Wasser) fehlt	4
Esch	Radfahren und spazieren über die Felder und an den Escher Seen möglich; Badegelegenheit; Wald fehlt; dadurch beschaulich, aber wenig ansprechend ('langweilig')	7
Zündorf	Hohe Erholungsqualität durch 'Freizeitinsel' Groov am Rhein; spazieren über Spargelfelder; Reiterhof; Schwimmbad; mit Personenfähre Ausflüge auf die gegenüberliegende Rheinseite	9
Kerpen	wie Sindorf, aber größere Nähe zum Wald; Naturschutzgebiet Kerpener Bruch hinter der Autobahn (über Brücke erreichbar, dennoch bildet die Autobahn eine Barriere)	6
Sindorf	Spazieren über Felder; Schwimmbad östlich von Sindorf; kein Park; besondere Attraktivität fehlt	5
Overath	Bewegte Topographie, attraktiver Wald und Flusstal (Agger) bieten Spazier- und Ausflugsmöglichkeiten	9
Heiligenhaus	ähnlich Overath, aber im unmittelbaren Umfeld eher Felder als Wald; Lage auf dem Berg bedeutet bei Spaziergängen: zuerst bergab, am Ende bergauf	8

**Tabelle 1: Erholungsqualität der Untersuchungsgebiete**

Quelle: Eigene Zusammenstellung nach Gespräch mit Birgit Kasper (vgl. Fußnote 8).

Um der unterschiedlichen Qualität von Grünflächen gerecht zu werden, wurde zur Bewertung der Erholungsqualität der Quartiere eine externe Expertin hinzugezogen<sup>8</sup>. In Tabelle 1 sind die der Bewertung zugrunde liegenden Einschätzungen zusammengefasst. Aufgrund dieser Einschätzungen wurde den Gebieten eine Erholungsqualität zugeordnet, die als relative Position der Gebiete zueinander auf einer Ordinalskala zu verstehen ist.

### 3.2.3 Schulen

Die Adressen der für die Untersuchungsgebiete relevanten Schulen wurden einer Veröffentlichung der Stadt Köln (2005) entnommen. Für die Gebiete außerhalb Kölns wurde die Lage der Schulen aus Stadtplänen und den Homepages der Städte entnommen. Dabei wurde kein festgelegter Radius um die Gebiete angenommen, sondern sichergestellt, dass für jede Schulart (Hauptschule, Realschule, Gymnasium) mindestens die nächstliegende Schule im Datensatz enthalten war. Den Adressen der Schulen wurden geographische Koordinaten zugeordnet. Grundschulen und Gesamtschulen wurden bei der Distanzberechnung ausgeschlossen, erstere weil die Stichprobe keine Kinder umfasst, Gesamtschulen weil diese lediglich in Form zweier privater Gesamtschulen in den Gebieten vorkommen.

## 4 Datenaufbereitung

Die vorliegenden Rohdaten erfordern für die Analyse umfangreiche Datenaufbereitungen. Diese werden im Folgenden für die einzelnen Bausteine des Projektansatzes beschrieben: Lebenslage, Lebensstil, Wohnstandortpräferenzen, Raumstruktur, Verkehrshandeln (inkl. Pkw-Verfügbarkeit). Davor wird kurz der Umgang mit fehlenden Werten dargestellt.

### 4.1 Behandlung fehlender Werte

Der Umgang mit fehlenden Werten spielt bei der Arbeit mit linearen Strukturgleichungsmodellen eine wichtige Rolle. Diese ist umso bedeutender, je höher die Anzahl der verwendeten Variablen ist, bei denen fehlende Werte auftreten. Gebräuchlich sind verschiedene Möglichkeiten, mit fehlenden Werten umzugehen:

- der Ausschluss von Fällen
- das Ersetzen fehlender Werte durch den Mittelwert der Variablen

<sup>8</sup> Mein Dank gilt hier Dipl.-Ing. Birgit Kasper, Büro für kommunale Entwicklungsplanung, Frankfurt am Main. Frau Kasper führte im Projekt StadtLeben ausführliche qualitative Interviews mit Bewohnern und Bewohnerinnen – teilweise auch mit lokalen Experten – der Quartiere. Darüber hinaus hat sie umfangreiche Begehungen und eine fotografische Dokumentation der Gebiete unternommen.



- die Imputation von Werten

Der Ausschluss von Fällen mit fehlenden Werten ist auf den ersten Blick ein sehr sauberes Verfahren, das häufig praktiziert wird. Mit zunehmender Anzahl von Variablen, bei denen fehlende Werte auftreten, reduziert sich allerdings die Anzahl verwertbarer Fälle. Dies ist hier ein durchaus nennenswertes Argument, denn obwohl die Stichprobe mit  $n=2.691$  Fällen groß genug ist, um auf einige Fälle zu verzichten, reduziert sie sich bei genauerem Hinsehen erheblich. Erstens wurden nur in sieben der zehn Untersuchungsgebiete Wohnstandortpräferenzen erfragt. Damit reduziert sich die Stichprobe auf  $n=2.133$ . Zweitens sind nicht alle Indikatoren der Verkehrsnachfrage für alle Befragten relevant. Besonders deutlich ist dies bei Berufswegen: Nur von 1.143 Personen liegen Angaben zur Wegelänge zur Arbeit vor, was in erster Linie daran liegt, dass nur 1.231 Personen zumindest in geringfügigem Umfang erwerbstätig sind. Drittens soll zur Validierung der Ergebnisse die Stichprobe in zwei unabhängige Teilstichproben geteilt werden (s.u.). Bei Berufswegen liegt die resultierende Stichprobe also bei rund  $n=570$ . Dies ist bei der Anwendung des asymptotisch verteilungsfreien Schätzverfahrens allenfalls gerade noch akzeptabel<sup>9</sup>.

Das Ersetzen fehlender Werte durch den Mittelwert der Variablen lässt lediglich die Mittelwerte, nicht aber die Verteilungen der Variablen unverändert. Da es hier vor allem auf die Varianzen und Kovarianzstrukturen ankommt, verbietet sich dieses Verfahren.

Für die Imputation von Werten stehen in LISREL zwei Verfahren zur Verfügung. Bei dem ersten handelt es sich um ein Matching-Verfahren. Dabei werden die fehlenden Werte als Mittelwerte der vergleichbaren Fälle geschätzt. Die Vergleichbarkeit wird über zu benennende Variablen definiert. Vermutet man beispielsweise, dass Geschlecht und Erwerbstätigkeit die wesentlichen Determinanten der Variable sind, deren Wert geschätzt werden soll, so würde bei einem teilzeitbeschäftigten Mann der Mittelwert aller teilzeitbeschäftigten Männer eingetragen. Dieses Verfahren hat allerdings den Nachteil, dass die Pfadkoeffizienten tendenziell überschätzt werden, weil bestehende Zusammenhänge zwischen Variablen auf Fälle mit fehlenden Angaben übertragen werden.

---

<sup>9</sup> Gegen den Ausschluss wechselnder Fälle je nach Modell sprechen auch praktische Gründe. AMOS kann bei bestimmten Operationen fehlende Werte nicht verarbeiten. Dies bedeutet, dass für jede einzelne Schätzung ein eigener Datensatz erstellt werden muss, in dem die jeweils relevanten Fälle mit fehlenden Werten gelöscht werden – es sei denn, man würde mit einem Datensatz arbeiten, in dem alle Fälle gelöscht sind, die in irgend einer Variablen, die in irgend einem Modell Verwendung findet, einen fehlenden Wert aufweisen. Dies ist allerdings hier aufgrund der Vielzahl relevanter Variablen nicht wünschenswert.

Darüber hinaus werden auch hier die Varianzen der Variablen verändert.

Beim zweiten Imputationsverfahren handelt es sich um die sogenannte Multiple Imputation mit EM-Algorithmus und Monte-Carlo-Markovkettensimulation. Dabei werden in iterativer Näherung (die anhand eines definierten Konvergenzkriteriums gestoppt wird) die fehlenden Werte unter Zuhilfenahme der empirischen Kovarianzmatrix geschätzt. Dieses Verfahren besitzt den Vorzug, dass über die Mittelwerte hinaus auch die Varianzen der Variablen unverändert bleiben. Deshalb wird hier dieses Verfahren angewandt. Die Imputation erfolgt in LISREL, weil ein ähnliches Verfahren in AMOS nicht zur Verfügung steht.

Die Imputation wird für die Merkmale der Lebenslage, des Lebensstils und der Wohnstandortpräferenzen vorgenommen. Retrospektive Angaben für die Zeit vor dem letzten Umzug werden nicht imputiert, weil sich dies kaum schätzen lässt. Der Vergleich der ursprünglichen Daten mit den imputierten Daten anhand der Lebenslage zeigt, dass Mittelwerte und Varianzen sich nicht verändert haben. Die Kovarianzen zwischen verschiedenen Merkmalen der Lebenslage ändern sich in einigen wenigen Fällen. Unter insgesamt 55 Korrelationen zwischen elf Variablen ergeben sich in drei Fällen Änderungen der Korrelation von 0.03 bis 0.05. Diese drei Fälle treten bei relativ starken Korrelationen im Zusammenhang mit dem Einkommen auf, wo recht viele fehlende Werte auftreten. So nimmt die Korrelation zwischen Äquivalenzeinkommen und Kindern im Haushalt von -0.46 auf -0.50 zu.

Beim Lebensstil und bei den Standortpräferenzen sind die Veränderungen der Korrelationen verschwindend gering. Sie liegen in der Regel unterhalb von 0.00, nur bei den Standortpräferenzen erreichen sie in einigen Fällen eine Größenordnung von 0.01.

Der Anteil fehlender Werte beträgt bei den Standortpräferenzen 2.9%, bei den Lebensstilen 0.1%, bei der Lebenslage 2.7% (meist beim Einkommen). Wenn alle Angaben zu einem Bereich fehlten, wurde der Fall gelöscht.

Insgesamt sind diese Veränderungen der Stichprobenmomente ohne Weiteres akzeptabel.

Beim Verkehrshandeln wurden keine fehlenden Werte imputiert. Bei fehlenden Werten der hier untersuchten Variablen handelt es sich in den meisten Fällen um vormalige Ausreißer, die bewusst gelöscht wurden, um die Linearität der Zusammenhänge, die Verteilung der Variablen und die Varianzaufklärung der Modelle zu verbessern. Bei einer probeweisen Imputation wurden die gelöschten Ausreißer gut reproduziert, was die Leistungsfähigkeit des Imputationsverfahrens zeigt.

	Originalvariablen mit Ausreißern*				ohne Ausreißer					
	N	Max	M	StD	K	N	Max	M	StD	K
Wochendistanz gesamt	2120	1759,9	130,4	160,5	<b>28,3</b>	2094	697,3	118,0	112,0	<b>4,2</b>
Wochendistanz Arbeit	2120	1750,0	56,4	116,4	<b>63,2</b>	2106	552,0	49,6	78,0	<b>8,1</b>
Wochendistanz Versorgung	2120	605,8	14,2	28,3	<b>253,8</b>	2066	60,7	11,3	10,2	<b>3,6</b>
Wochendistanz Freizeit	2120	1413,2	59,6	88,3	<b>57,2</b>	2075	298,6	50,3	52,8	<b>4,1</b>
Wochendistanz MIV	2120	1750,0	93,2	138,3	<b>35,7</b>	2095	585,4	83,2	97,4	<b>3,8</b>
Distanz je Aktivität gesamt	2120	335,7	10,7	14,6	<b>138,2</b>	2111	96,3	10,0	10,1	<b>11,5</b>
Distanz je Aktivität Freizeit	2118	612,6	13,0	25,9	<b>186,6</b>	2103	217,8	12,3	19,2	<b>36,6</b>

**Tabelle 2: Vergleich der Stichproben (Umfang, Verteilung, Stichprobenmomente) mit und ohne Ausreißer**

N=Stichprobenumfang (netto), Max=Maximum, M=Mittelwert, StD=Standardabweichung, K=Kurtosis, Standardfehler der Kurtosis jeweils ca. 0.1. Minimum aller Variablen 0

\* n=2133 inkl. 13 Fälle ohne Angaben zu Wohnstandortpräferenzen. Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Bei einigen Variablen des Verkehrshandelns wurden Ausreißer eliminiert, um die Modellanpassung zu verbessern. Die Entscheidung darüber wurde jeweils nach umfangreichen Regressionsanalysen getroffen, bei denen Schritt für Schritt Ausreißer ausgeschlossen wurden, bis keine wesentliche Verbesserung mehr eintrat. Tabelle 2 zeigt die Stichprobenumfänge und Stichprobenmomente vor und nach der Eliminierung von Ausreißern. Es wird deutlich, dass die Standardabweichungen und insbesondere die Kurtosis der Variablen stark verringert werden. Auch die Varianzaufklärung wird deutlich verbessert. So beträgt nach den erwähnten Analysen die Varianzaufklärung der wöchentlichen Pkw-Reisedistanz ohne Ausreißer  $R^2=0.36$ , mit Ausreißern  $R^2=0.24$ . Bei der wöchentlichen Reisedistanz bei Versorgungswegen betragen die Werte  $R^2=0.09$  (ohne Ausreißer) bzw.  $R^2=0.03$  (mit Ausreißern).

Damit stehen in Abhängigkeit vom verwendeten Indikator des Verkehrshandelns maximal 2.120 Fälle für die Analyse zur Verfügung.

## 4.2 Lebenslage I

Die Lebenslage der Befragten ist in den StadtLeben-Daten durch ein differenziertes Set von demografischen und sozialstrukturellen Merkmalen erhoben worden. Bedingt durch den Projektablauf erfolgte die Erhebung nicht in allen Untersuchungsgebieten auf exakt die gleiche Art und Weise. Durch geeignete Umcodierungen lassen sich jedoch vergleichbare Variablen herstellen. Die relevanten Merkmale umfassen die Haushaltsstruktur (Anzahl der Kinder bis 16 Jahre, Haushaltstyp, Haushaltsgröße), die Stellung im Erwerbsprozess (ggf. mit dem Umfang der Erwerbstätigkeit), die Stellung im Beruf, den höchsten Schulabschluss sowie Haushaltseinkommen, Geschlecht, Alter und Nationalität.

Zum großen Teil handelt es sich dabei um nominal- und ordinalskalierte Variablen, die in linearen Strukturgleichungsmodellen (sofern die ‚klassische‘ ML-Schätzmethode angewendet wird) nicht einsetzbar sind, es sei denn, als exogene Variablen. Zudem

würde die große Zahl der relevanten Variablen die Modelle leicht unüberschaubar machen.

Deshalb werden die Variablen faktorenanalytisch auf eine handhabbare Zahl von Dimensionen reduziert, was sich insofern ohnehin anbietet, weil viele Lebenslagemerkmale stark miteinander korrelieren. Zuvor werden insbesondere die ordinalskalierten Variablen durch geeignete Transformationen in metrische Variablen umcodiert. Von Fall zu Fall können auch metrisch skalierte Originalvariablen in den Modellen verwendet werden. Die Arbeit mit Faktorwerten als ‚gemessenen Variablen‘ unter Ausklammerung von Messmodellen ist in der Verkehrsforschung nicht unüblich (z.B. Cao/Mokhtarian/Handy 2007, Scheiner/Holz-Rau 2007), weil Modelle mit latenten Variablen häufig zu äußerst geringer Modellgüte führen (Simma 2000).

Tabelle 3 zeigt die in der Faktorenanalyse verwendeten Variablen sowie das Ergebnis. Die Faktoren erklären 58.1% der Varianz der Eingangsvariablen. Dies ist als durchaus akzeptabel anzusehen. Sensitivitätsanalysen zeigen, dass die Faktorenstruktur stabil ist<sup>10</sup>.

<sup>10</sup> Bei einigen Variablen wurden verschiedene Variationen verglichen. So fand etwa statt der Anzahl der Kinder im Haushalt als Variante eine binäre Variable Verwendung (Kinder ja/nein). Statt des Umfangs der Erwerbstätigkeit wurde als Variante eine Binärvariable „Vollzeit/Teilzeit Erwerbstätig“ (ja/nein) verwendet. Anstelle der Binärvariablen „mindestens zwei Erwachsene im Haushalt“ wurde probeweise auch die Anzahl der Erwachsenen im Haushalt verwendet. Dabei wird das Ergebnis allerdings dadurch verzerrt, dass Wohngemeinschaften eine höhere Ausprägung haben als Paarhaushalte, obwohl ihre einzelnen Mitglieder im Sinne des Lebens- und Familienzyklus eher den Singles ‚nahe stehen‘. Die Nationalität wurde ausgeklammert, da sie stark mit dem demografischen Faktor ‚Familie‘ korrelierte und ihn damit zu einem Faktor ‚ausländische Familie‘ machte, der im verkehrlichen Sinne kaum zu interpretieren ist, sofern ihm kein Faktor ‚deutsche Familie‘ zum Vergleich gegenübersteht.

	junge Erwerbs- tätige / Status	Fami- lie	Männ- liche Führungs- personen
Umfang Erwerbstätigkeit (Anzahl Wochenstd) <sup>1</sup>	0,761	-0,114	0,208
Alter	-0,735	-0,342	0,284
Schulbildung (Dauer in Jahren) <sup>2</sup>	0,632	-0,068	0,249
Äquivalenzeinkommen <sup>3</sup>	0,171	-0,742	0,270
Anzahl der Kinder unter 16 Jahre im Haushalt	0,267	0,733	-0,099
Mindestens zwei Er- wachsene im Haushalt?	-0,178	0,657	0,474
Geschlecht weiblich	-0,033	0,082	-0,709
Leitende Position im Beruf <sup>4</sup>	0,273	-0,067	0,513

**Tabelle 3: Dimensionen der Lebenslage**

<sup>1</sup> Schätzung aus kategorialen Angaben: Vollzeit Erwerbstätig = 40 Std. (Stunden/Woche), Teilzeit = 20 Std.; geringfügig beschäftigt = 12 Std.; unregelmäßige kleinere Jobs = 8 Std.; nicht erwerbstätig = 0 Std.

<sup>2</sup> Schätzung aus kategorialen Angaben: kein Abschluss = 8 Jahre; Volks-/Hauptschule = 9 Jahre; Mittlere Reife = 10 Jahre; (Fach-) Abitur = 13 Jahre; (Fach-) Hochschulabschluss = 18 Jahre. Für Befragte, die sich noch in der Schulausbildung befinden, wurden aus den entsprechenden Angaben sowie dem Alter Werte zwischen 10 und 13 geschätzt.

<sup>3</sup> Haushaltseinkommen bezogen auf die Zahl der Haushaltsmitglieder, wobei aufgrund der mit der Haushaltsgröße sinkenden Lebenshaltungskosten pro Kopf jedes weitere Haushaltsmitglied über die erste Person hinaus mit 0,8 gewichtet wurde.

<sup>4</sup> Beamte im höheren oder gehobenen Dienst, leitende Angestellte, Selbstständige und Freiberufler.

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Der erste Faktor kann als „junge Erwerbstätige“ interpretiert werden. Durch die hohe Ladung der Schulbildung besitzt er eine starke Status-Komponente. Eine negative Ausprägung dieses Faktors besitzen in erster Linie Rentner.

Der zweite Faktor lässt sich als „Familie“ beschreiben. Er zeichnet sich vor allem durch demografische Merkmale aus: Kinderanzahl, mit Partner lebend, eher junges Alter (schwächere Ladung). Aber auch das Einkommen lädt stark auf diesem Faktor, und zwar im negativen Sinne, weil das Pro-Kopf-Einkommen in größeren Haushalten in der Regel niedriger ist.

Der dritte Faktor ist durch den sozialen Status gekennzeichnet und lässt sich als „männliche Führungspersonen“ bezeichnen. Er ist durch eine negative Ladung des Geschlechts gekennzeichnet (hohe Ausprägung bei Männern), eine positive Ladung der Variablen „leitende Position im Beruf“ sowie ergänzend durch das Zusammenleben mit einem Partner bzw. in diesem Falle eher einer Partnerin. Nach den

mäßigen Ladungen anderer Variablen gehen in diesen Faktor außerdem mit geringerem Gewicht hohes Einkommen, hohe Schulbildung, Erwerbstätigkeit und überdurchschnittliches Alter ein. Hohe Ausprägungen auf diesem Faktor dürften also vor allem Männer im gesetzteren Alter in hoher Position besitzen, die entweder keine oder bereits ältere/erwachsene Kinder haben.

Insgesamt ergaben sich durch die Analyse drei gut interpretierbare und differenzierte Faktoren, mit denen in den Strukturgleichungsmodellen gearbeitet werden kann. Alternativ können mit nicht-parametrischen Schätzverfahren auch originalskalierte Variablen der Lebenslage verwendet werden.

### 4.3 Lebenslage II

Die Arbeit mit Faktorwerten als 'gemessenen Variablen' ist allerdings theoretisch unbefriedigend, weil der Ausschluss der Messmodelle im Grunde der Annahme entspricht, dass die verwendeten Variablen fehlerfreie Messungen der interessierenden latenten Konstrukte sind. Bei der weiteren Arbeit wurden deshalb bei der Abbildung der Lebenslage in den Modellen auch Messfehler zugelassen, d.h. die Lebenslage wurde in Form konfirmatorischer Faktorenanalysen als Messmodell in den Modellen abgebildet. Dabei wurden verschiedene Varianten unternommen, weil sich zeigte, dass die Lebenslage-Variablen sich nicht ohne Weiteres in einem befriedigenden Messmodell zusammenfassen ließen. Folgende Variablen wurden berücksichtigt:

- Umfang Erwerbstätigkeit (Anzahl Wochenstunden)
- Alter
- Schulbildung (Dauer in Jahren)
- Äquivalenzeinkommen
- Anzahl der Kinder unter 16 Jahre im Haushalt
- Geschlecht
- Haushaltsgröße.

Weitere Variablen wurden ausgeklammert, weil sie nur in kategorialer Form vorlagen (z.B. leitende Position im Beruf ja/nein). Es wurden sowohl Analysen mit den originalen als auch mit standardisierten Variablen unternommen. Das Alter wurde probeweise auch in Form einer nicht-linearen Transformation berücksichtigt, bei der die Ausprägungen bis zur Lebensmitte (40 Jahre) zu-, danach aber wieder abnahmen. Dadurch wurde die Korrelationen mit der Anzahl der Kinder im Haushalt, aber auch mit dem Umfang der Erwerbstätigkeit höher. Generell wurde so vorgegangen, dass zunächst ein möglichst einfaches und überschaubares Modell die Grundlage bildete. Auf der Basis der Werte für die Modellanpassung, der Residuen und der Modifikationsindizes wurde dann dieses Modell iterativ induktiv wei-

terentwickelt. Die Modifikationsindizes geben die Verbesserung der Modellanpassung ( $\chi^2$ -Wert) wider, die mit einer Freisetzung des entsprechenden Pfades bzw. der Kovarianz erzielt würde. Als generelle Ergebnisse dieser Vorgehensweise sind festzuhalten:

- Es gelingt nicht, die Variablen der Lebenslage in mehreren scharf voneinander zu trennenden latenten Variablen abzubilden. Dies gilt unabhängig davon, ob die latenten Variablen als voneinander unabhängige Dimensionen modelliert werden oder ob Korrelationen zwischen ihnen zugelassen werden.
- Als am 'sperrigsten' erweist sich dabei das Geschlecht, das sich keiner latenten Variable zuordnen lässt. Geht man davon aus, dass das Geschlecht am ehesten eine latente Variable 'Gender' abbildet, dann wäre 'Gender' neben dem Geschlecht vermutlich am ehesten durch weitere Selbstauskünfte über ausgeübte Geschlechterrollen abbildbar (z.B. Rollenverteilung im Haushalt, evtl. auch über hier als Teil des Lebensstils betrachtete Items wie die Freizeitpräferenz 'mit Kindern spielen'), jedenfalls aber nicht durch die hier erhobenen 'harten' Lebenslagemerkmale.
- Die anfänglich unzureichende Modellanpassung lässt sich durch die Zulassung von Fehlerkovarianzen und die weniger scharfe Trennung der Faktoren (Zulassung von Pfaden von einer manifesten Variablen zu mehreren latenten Variablen) deutlich verbessern. Gleichzeitig treten dabei aber immer wieder negative Varianzschätzungen auf.

Eine rundum befriedigende Lösung zu finden, ist demnach praktisch nicht möglich. Es gilt stattdessen einen Kompromiss zu finden zwischen theoretischer Tragfähigkeit des Messmodells (inhaltliche Interpretierbarkeit), akzeptabler Modellanpassung und Lösbarkeit des Gleichungssystems.

Die letztlich präferierte Lösung sieht zwei latente Variablen 'sozialer Status' und 'Familie' vor (Abbildung 2). Das Geschlecht wird als Einzelvariable verwendet und ist nicht Teil des Messmodells. Die beiden latenten Variablen sind nicht trennscharf. Die latente Variable 'sozialer Status' wird abgebildet durch drei manifeste Variablen: Schulbildung, Umfang der Erwerbstätigkeit und Äquivalenzeinkommen. Die latente Variable 'Familie' wird durch vier manifeste Variablen abgebildet: Anzahl der Kinder unter 16 Jahre im Haushalt, Haushaltsgröße, Alter und Äquivalenzeinkommen. Darüber hinaus werden eine Reihe von Fehlerkovarianzen zwischen den gemessenen Variablen freigegeben. Dies wird in den vollständigen Modellen von Fall zu Fall entschieden (vgl. Kapitel 6.3).

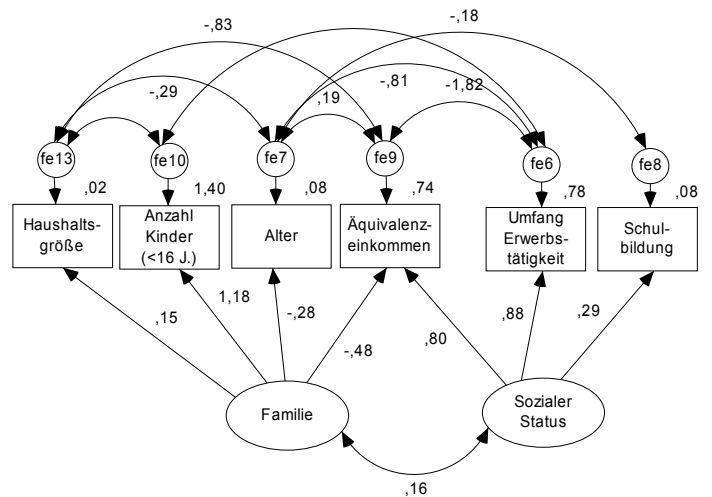


Abbildung 2: Messmodell der Lebenslage

CMIN/DF=6.046, RMSEA=0.056, Hoelter  
(p=0.05)=1023.

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

## 4.4 Lebensstil I

### 4.4.1 Überblick über die Vorgehensweise

'Lebensstil' hat sich in der Forschung zu einem schillernden Begriff entwickelt, der auf sehr unterschiedliche Art definiert und operationalisiert wird. Nach Spellerberg (1996:57) sind Lebensstile "gruppenspezifische Formen der Alltagsorganisation und -gestaltung, die auf der Ebene des kulturellen Geschmacks und der Freizeitaktivitäten symbolisch zum Ausdruck kommen". Mit Müller (1992) lassen sich vier Dimensionen von Lebensstilen unterscheiden:

- expressive Dimension (z.B. Freizeitpräferenzen/-verhalten, Alltagsästhetik, Konsum)
- interaktive Dimension (z.B. soziale Kontakte, Kommunikation)
- evaluative Dimension (z.B. Werte, Lebensziele, Wahrnehmungen)
- kognitive Dimension (z.B. Selbstidentifikation, Zugehörigkeit).

In den StadtLeben-Daten sind Lebensstile anhand von insgesamt 35 Items erfasst, die auf drei Bereiche abzielen: Freizeitpräferenzen, Lebensziele und Wertorientierungen, und alltagsästhetische Schemata (kultureller Geschmack). Ergänzend wurde der Umfang des sozialen Netzwerkes mit aufgenommen. Dem liegt im Wesentlichen der Lebensstil-Ansatz des Wissenschaftszentrums Berlin für Sozialforschung zugrunde (Spellerberg 1996). Die erfragten Items lassen sich aber auch in das erwähnte Schema von Müller (1992) einordnen.

Lebensstile spielen in den Strukturgleichungsmodellen die Rolle intervenierender Variablen und sollen durch metrische Variablen abgebildet werden. Dafür

bietet es sich an, nicht die Zugehörigkeit zu einer Lebensstilgruppe als Variable zu verwenden – also eine Typisierung der Befragten –, sondern die Ausprägung auf einer oder mehreren Lebensstilskalen. Dies ist auch inhaltlich plausibler als die Verwendung von Lebensstiltypen, da eher lineare als diskrete Zusammenhänge zwischen Lebensstil und Verkehrshandeln zu erwarten sind<sup>11</sup>.

Aus den vorliegenden Items lassen sich nun auf mehrere Arten Skalen bilden. Zum einen können die Items mit Hilfe von Faktorenanalysen auf grundlegende, latente (nicht direkt erfragte) Dimensionen verdichtet werden. Basierend auf theoretischen Überlegungen können aber zum anderen auch einige interessierende Items ausgewählt und zu Skalen oder Faktoren zusammengefasst werden.

Im Projekt StadtLeben wurden mit Faktorenanalysen acht Lebensstil-Indikatoren („Dimensionen“) ermittelt. Bei der Verwendung dieser Dimensionen sind allerdings mehrere Dinge zu bedenken:

1. Die acht Lebensstil-Indikatoren sind keine echten Dimensionen, weil die Faktorenanalysen getrennt für die drei Bereiche Freizeitpräferenzen, Lebensziele und Wertorientierungen, und alltagsästhetische Schemata durchgeführt wurden. Die ermittelten Indikatoren sind also nicht unabhängig voneinander.
2. Anhand der (auf der Basis der Lebensstil-Indikatoren) clusteranalytisch ermittelten Lebensstilgruppen wurde deutlich, dass es sich möglicherweise bei den Gruppen teilweise um Artefakte handelt, die durch eine generell über- oder unterdurchschnittliche Zustimmungsbereitschaft einiger Befragter entstehen. So zeichnete sich die Gruppe der „Distanzierten“ durch eine generelle Zurückhaltung aus, die Gruppe der „Außerhäuslich-Geselligen“ durch eine generelle Zustimmungstendenz zu den Items.
3. Die Varianzaufklärung der Faktoren ist teilweise sehr mäßig, bedingt durch geringe Korrelationen zwischen vielen Items.

Aufgrund dieser Überlegungen wurden hier Lebensstil-Indikatoren neu gebildet. Dabei wurde ein schrittweises Vorgehen angewandt.

Zunächst wurden die Items mit Hilfe von Faktorenanalysen über alle Items hinweg neu verdichtet. Um dem oben genannten zweiten Einwand Rechnung zu

tragen, wurden die Items vorher transformiert. Die Ausprägungen der Items wurden als Abweichungen vom Mittelwert der Ausprägungen aller Lebensstilitems der jeweiligen Person umcodiert, um die generelle tendenzielle Zustimmung- oder Ablehnungstendenz einiger Personen zu eliminieren.

Daraufhin wurden umfangreiche Regressionsanalysen des Verkehrshandelns mit den ermittelten Dimensionen unternommen. Die Erwartung war, dass nur einige wenige Dimensionen des Lebensstils für die Verkehrsnachfrage von großer Bedeutung sind, etwa eine starke außerhäusliche Freizeitorientierung oder eine starke Karriereorientierung. Dagegen war die Erwartung, dass Lebensstildimensionen, die sich vorwiegend auf den ästhetischen Geschmack beziehen, für die Verkehrsnachfrage bedeutungslos seien. Damit hätten die relevanten Lebensstildimensionen auf eine geringe Anzahl reduziert werden können. Diese Erwartung erfüllte sich jedoch nicht. Stattdessen zeigten sich eher disparate und schwer interpretierbare Zusammenhänge zwischen einzelnen Lebensstildimensionen und den verschiedenen Merkmalen des Verkehrshandelns.

Aufgrund dieses Ergebnisses, das im Wesentlichen auf einem datengeleiteten Vorgehen beruhte, wurde die Vorgehensweise verändert. Anhand theoretischer bzw. an Plausibilität orientierter Überlegungen wurden die berücksichtigten Items stark reduziert und in mehreren separaten Faktorenanalysen zu Lebensstilindikatoren verdichtet, von denen ein Einfluss auf bestimmte Aspekte des Verkehrshandelns zu erwarten ist. Dieses Vorgehen hat den Vorzug, dass es aufgrund seiner theoretischen Orientierung eher reproduzierbar ist und damit die Ergebnisse überprüfbar sind. Dagegen ist das empiristische Vorgehen stärker den Zufälligkeiten eines Datensatzes ausgesetzt.

Die genaue Vorgehensweise wird in Kapitel 0 beschrieben. Im folgenden Kapitel 4.4.2 wird aber zunächst die schließlich verworfene Vorgehensweise dokumentiert.

#### 4.4.2 Schritt 1: Faktorenanalytische Bildung von Lebensstildimensionen

Wie oben beschrieben, wurden zunächst unabhängige Lebensstildimensionen faktoranalytisch gebildet. Dabei wurde eine Faktorenanalyse über alle erfragten Lebensstil-Items durchgeführt. Der im Projekt StadtLeben ursprünglich ebenfalls als Lebensstilaspekt betrachtete Umfang des sozialen Netzwerks wurde aus Gründen der Skalierung ausgeklammert. Die Ausprägungen der Items wurden als Abweichungen vom Mittelwert der Ausprägungen aller Lebensstilitems der jeweiligen Person umcodiert, um die generelle tendenzielle Zustimmung- oder Ablehnungstendenz einiger Personen zu eliminieren.

<sup>11</sup> Beispielsweise ist der Zusammenhang "je stärker die außerhäusliche Freizeitorientierung, desto mehr Freizeitwege" plausibler als "wenn starke außerhäusliche Freizeitorientierung, dann viele Freizeitwege", weil eine "starke außerhäusliche Freizeitorientierung" nicht in diskreter Form entweder zutrifft oder nicht zutrifft. Vielmehr bewegt sich die außerhäusliche Freizeitorientierung auf einer stetigen Skala.

	Trivialschema	Spannungsschema	Anerkennung + Geborgenheit	Familie	Unternehmungslust	Information*	Hochkulturschema (lesen)*	Selbstverwirklichung, Verantwortung*	Kreativität
TV: Unterhaltungsserien	0,69	-0,06	0,01	-0,05	0,07	0,18	0,08	-0,01	-0,01
TV: Shows, Quiz	0,66	-0,03	0,07	-0,03	-0,08	0,01	0,14	0,04	-0,07
Frz: Fernsehen, Video	0,62	0,16	-0,02	-0,02	-0,12	-0,21	0,24	0,11	0,02
TV: Heimatfilme	0,54	-0,04	0,13	0,04	-0,36	0,06	-0,15	0,11	0,09
Lesen: Arzt-/Schicksalsromane	0,45	0,06	0,11	-0,06	-0,18	0,22	-0,17	0,21	-0,26
TV: Horror	0,05	0,71	0,03	-0,13	0,06	0,13	0,12	0,00	0,11
TV: Action	0,09	0,70	-0,01	-0,02	-0,03	0,13	0,31	-0,04	0,02
Lesen: Comics	-0,14	0,65	-0,16	-0,07	0,02	0,04	-0,03	-0,01	-0,13
Ziele: gut aussehen	0,14	-0,06	0,67	-0,12	0,03	0,10	0,04	0,04	0,07
Ziele: Anerkennung durch andere	-0,04	-0,03	0,60	-0,16	-0,08	0,15	0,13	-0,17	-0,16
Ziele: Sicherheit, Geborgenheit	0,07	-0,08	0,49	0,20	-0,26	-0,08	0,11	0,33	0,00
Ziele: für andere da sein	0,01	0,01	0,43	0,36	0,10	-0,06	-0,22	0,10	0,32
Frz: Computer	-0,27	0,24	-0,33	-0,14	0,11	0,20	0,29	-0,22	0,06
Frz: mit Familie beschäftigen	-0,03	-0,11	-0,03	0,78	-0,02	0,03	0,08	0,07	-0,04
Frz: mit Kindern beschäftigen	-0,06	-0,11	-0,13	0,75	-0,02	0,19	0,09	-0,04	-0,08
Frz: private Treffen	-0,01	-0,03	0,03	0,25	0,66	0,07	0,11	0,14	0,06
Frz: Kino, Theater, Konzert	-0,30	-0,01	-0,05	-0,29	0,60	0,02	0,02	0,07	-0,03
Ziele: Sparsamkeit	0,15	-0,03	0,30	0,00	-0,51	-0,03	0,05	0,25	0,05
Ziele: aufregendes, abwechslungs- reiches Leben	-0,07	0,17	0,30	-0,15	0,40	0,14	0,01	-0,32	0,11
Frz: Gartenarbeit	-0,22	-0,33	-0,13	0,17	-0,35	-0,01	0,29	0,24	-0,01
TV: Dokumentationen	-0,12	-0,13	-0,13	-0,13	-0,11	-0,74	-0,02	0,05	0,07
TV: politisch/kulturelle Magazine	-0,11	-0,29	-0,09	-0,13	-0,02	-0,73	-0,03	-0,12	-0,03
Frz: Kurse, Weiterbildung	-0,32	-0,27	-0,23	-0,27	0,10	0,33	0,15	-0,13	-0,11
Lesen: Gedichte	-0,10	-0,15	-0,05	-0,07	-0,08	0,06	-0,73	0,01	0,03
Lesen: Literatur	-0,34	-0,17	-0,19	-0,13	0,01	-0,15	-0,62	0,01	-0,14
Ziele: Führungspositionen	-0,15	0,04	0,01	-0,03	0,02	-0,02	0,23	-0,69	0,00
Ziele: Politisches, gesellschaftliches Engagement	-0,29	-0,18	-0,06	-0,01	-0,01	-0,16	-0,17	-0,51	-0,02
Frz: Spazieren, Wandern, Ausflug	-0,22	-0,29	0,08	0,00	0,00	-0,11	0,17	0,41	-0,03
Lesen: Unterhaltungsromane	0,05	-0,03	-0,12	-0,03	0,18	0,14	-0,24	0,26	-0,68
Frz: Musik hören	0,02	0,04	-0,08	-0,21	0,26	0,08	-0,12	0,28	0,52
Frz: künstlerische Tätigkeiten	-0,21	-0,14	-0,31	-0,11	-0,04	0,38	-0,17	0,03	0,40

**Tabelle 4: Lebensstildimensionen – Matrix der Faktorladungen**

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

	Arbeits- wege	Versorgungs- wege	Freizeit- wege	MIV- Nutzung	ÖV- Nutzung	NMIV- Nutzung
Frz: Fernsehen, Video	-(x)		-(x)			
Frz: Spazieren, Wandern, Ausflug			(x)			
Frz: Musik hören	-(x)		-(x)			
Frz: künstlerische Tätigkeiten			(x)			
Frz: Computer			-(x)			
Frz: mit Familie beschäftigen		x				
Frz: mit Kindern beschäftigen		x				
Frz: Gartenarbeit	-(x)	(x)				
Frz: Kurse, Weiterbildung	(x)		x			
Frz: private Treffen	-(x)		(x)			
Frz: Kino, Theater, Konzert			x			
Ziele: gut aussehen						
Ziele: Anerkennung durch andere	(x)					
Ziele: Sicherheit, Geborgenheit						x
Ziele: für andere da sein						(x)
Ziele: Sparsamkeit						x
Ziele: aufregendes, abwechslungsreiches Leben	(x)		x	(x)	(x)	
Ziele: Führungspositionen	x			x	x	
Ziele: Politisches, gesellschaftliches Engagement	x			x	x	
TV: Heimatfilme						(x)
TV: Horror						
TV: Action						
TV: Unterhaltungsserien						(x)
TV: Shows, Quiz						(x)
TV: Dokumentationen	(x)					
TV: politisch/kulturelle Magazine	(x)					
Lesen: Comics						
Lesen: Arzt-/Schicksalsromane						(x)
Lesen: Gedichte						
Lesen: Literatur						
Lesen: Unterhaltungsromane						(x)

**Tabelle 5: Hypothetisch relevante Lebensstil-Items für verschiedene Aspekte des Verkehrshandelns**

x wichtiges Item, (x) weiteres mögliches Item, -(x) weiteres mögliches Item mit negativer Korrelation

Nach dem Kaiser-Kriterium (Eigenwert > 1) ergeben sich zehn Faktoren, von denen allerdings einer aus nur einem Item besteht (Ablehnung von Unterhaltungsliteratur). Die Zahl der Faktoren wurde deshalb im nächsten Schritt auf neun reduziert. Nach mehreren Varianten wurde die Hauptkomponentenmethode mit Varimax-Rotation angewandt. Die aufgeklärte Varianz ist mit 52,5% mäßig, aber akzeptabel. Die Faktoren sind inhaltlich interpretierbar, aber nicht überall trennscharf.

Probeweise getrennte Faktorenanalysen über die zwei Bereiche Freizeit sowie Lebensziele/Werte/Alltagsästhetik führen zu deutlich klarerer Interpretierbarkeit, aber ebenfalls mäßigen Varianzaufklärungs-raten.

Gruppenbildungen anhand von Clusteranalysen zeigen darüber hinaus, dass die Zuordnung zu Gruppen nur sehr unscharf möglich ist. In diversen Varianten mit vier bis sechs Gruppen liegt jeweils nur etwa 16% bis 22% der Gesamtvarianz der Dimensionen zwischen den Gruppen. Die Cluster sind also intern äußerst heterogen. Damit sei lediglich angedeutet, wie stark die Vereinfachung ausfällt, die durch die Zuordnung von Personen zu Lebensstilgruppen anhand von verschiedenen Dimensionen in Kauf genommen wird.

#### 4.4.3 Schritt 2: Bildung von jeweils an ein Modell inhaltlich angepassten Faktoren

Aufgrund der oben erwähnten Probleme wurde schließlich ein anderes Vorgehen gewählt. In den Strukturgleichungsmodellen ist die Unabhängigkeit

verschiedener Lebensstildimensionen untereinander nicht erforderlich. Ohnehin sollen Lebensstile in den Modellen so einfach wie möglich abgebildet werden, möglichst auf der Grundlage von nur einer Dimension. Es ist aber auch möglich, mehrere miteinander korrelierte Lebensstilvariablen in den Modellen zu verwenden.

Deshalb wird im Folgenden eine an verschiedene Indikatoren des Verkehrshandelns inhaltlich angepasste Bildung von Lebensstilindikatoren vorgenommen. So sind beispielsweise vermutlich für die Häufigkeit und Länge von Freizeitwegen andere Aspekte des Lebensstils maßgeblich als für die Nutzung des MIV. Zur Bildung der Indikatoren werden zwei Methoden angewandt: Erstens wird die Verlässlichkeit der Item-Kombinationen mit Reliabilitätsanalysen auf der Grundlage von Cronbach's  $\alpha$  geprüft, zweitens werden Faktorenanalysen mit den jeweils ausgewählten Items durchgeführt. Die Items werden gezielt so ausgewählt, dass die entstehenden Faktoren erstens ein akzeptables Maß an Varianz der Original-Items aufklären und zweitens möglichst genau die interessierenden Indikatoren abbilden. Ziel ist also weder die möglichst umfassende Abbildung des Lebensstils noch die Unabhängigkeit aller Lebensstilindikatoren (im Sinne von Dimensionen oder Faktoren) untereinander. Das Vorgehen wird im Folgenden für verschiedene Aspekte des Verkehrshandelns ausgeführt.

#### 4.4.3.1 Berufswege

Für Berufswege wird vor allem die Karriereorientierung als relevant angesehen. Diese drückt sich vor allem in zwei der erhobenen Items aus: die Lebensziele ‚Führungspositionen‘ und ‚politisches oder gesellschaftliches Engagement‘. Beide sind positiv miteinander korreliert. Weitere Items, die dazu in Bezug stehen könnten, sind das Interesse an politischen und kulturellen Magazinen sowie an Dokumentationen, beide als Ausdruck für eine sachliche, politisch-gesellschaftliche Orientierung. Des Weiteren könnten die Lebensziele ‚ein aufregendes, abwechslungsreiches Leben führen‘ und ‚Anerkennung durch andere‘ sowie die Häufigkeit von Kursen und Weiterbildung in der Freizeit zur beruflichen bzw. Karriere-Orientierung in Bezug stehen. Im negativen Sinne könnten damit eher passive, häusliche oder zeitintensive Freizeitaktivitäten korrelieren, etwa Fernsehen, Musik hören, Gartenarbeit oder private Treffen.

Im Ergebnis zeigt sich allerdings, dass die Berücksichtigung weiterer Items neben ‚Führungsposition‘ und ‚politisches oder gesellschaftliches Engagement‘ die Skalierung nur verschlechtert. Deshalb wird ausschließlich auf diese beiden Items zurückgegriffen. In der Reliabilitätsanalyse führt dies zu dem Problem, dass Cronbach's  $\alpha$  auch bei eher schwachen Korrelationen mit der Anzahl der berücksichtigten Items zunimmt, so dass der Wert hier auf der Grundlage von nur zwei Items eher niedrig ist ( $\alpha=0.48$ ).

Eine Faktorenanalyse führt aber zu einem guten Ergebnis. Sie lässt sich gemeinsam mit zwei weiteren Items durchführen, die für eine traditionelle Wertorientierung stehen und die für die Nutzung des NMIV als relevant angesehen werden (s.u.). Die vier Items lassen sich zu zwei Faktoren zusammenfassen. Die aufgeklärte Varianz beträgt 66.4%. Die Ladungen sind in Tabelle 6 dargestellt. Der hier relevante Faktor wird als Wertorientierung ‚Selbstverwirklichung‘ interpretiert und beruht im Schwerpunkt auf einer beruflichen (Karriere-)Orientierung.

	Traditionell	Selbstverwirklichung
Sparsamkeit	<b>0,814</b>	-0,030
Sicherheit, Geborgenheit	<b>0,810</b>	-0,013
politisches, gesellschaftliches Engagement	0,070	<b>0,819</b>
Führungsposition übernehmen	-0,115	<b>0,805</b>

**Tabelle 6: Lebensstil: Dimensionen von Wertorientierungen – Matrix der Faktorladungen**

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

#### 4.4.3.2 Freizeitwege

Bei Freizeitwegen sind die größten Lebensstileinflüsse zu erwarten. Dabei können verschiedene Aspekte des Lebensstils von Bedeutung sein: eine häusliche Orientierung, eine außerhäusliche Orientierung, eine Selbstverwirklichungsorientierung oder eine traditionelle Orientierung. Im Projekt StadtLeben wurden auf der Grundlage der hier verwendeten Daten acht Lebensstilindikatoren gebildet: Wertorientierung ‚traditionell‘, Wertorientierung ‚Selbstverwirklichung‘, außerhäusliche Freizeitorientierung, häuslich-familiäre Freizeitorientierung, Trivialekultur, Hochkultur (lesen), Hochkultur (fernsehen), Spannungskultur (Schweer/Hunecke 2006). Alle diese Dimensionen können für bestimmte Freizeitaktivitäten hohe Bedeutung besitzen. Dies hängt vor allem von der Differenzierung der untersuchten Freizeitaktivitäten ab. So wird beispielsweise eine starke hochkulturelle Orientierung in starkem Maß die Häufigkeit der Besuche klassischer Konzerte beeinflussen, weniger aber die Häufigkeit von Gastronomiebesuchen (sofern nicht unterschieden wird, welche Art gastronomischer Einrichtung gemeint ist).

Für die gemeinsame Untersuchung aller Freizeitaktivitäten werden hier vor allem die Freizeitpräferenzen als wichtig angesehen. Allerdings klärt die aus dem Projekt StadtLeben vorliegende Faktorenstruktur nur einen geringen Anteil der Gesamtvarianz der Freizeit-Items auf (34.4%). Dies liegt an den teilweise eher geringen Korrelationen zwischen den Freizeitpräferenzen<sup>12</sup>. Deshalb wurden in einem stufenwei-

<sup>12</sup> Eine Faktorenanalyse auf der Grundlage des Kaiser-Kriteriums (Eigenwert > 1) führt zu einer Extraktion von neun (!) Faktoren aus elf Items.



sen Vorgehen nach und nach einige Items aus der Faktorenanalyse ausgeschlossen, bis sich eine vergleichbare Struktur mit zwei Faktoren, aber akzeptabler Varianzaufklärung (53,0%) fand (Tabelle 7).

	Außer-häuslich	Häuslich-Familiär
Kurse, Weiterbildung	<b>0,709</b>	0,089
Kino, Theater, Konzert	<b>0,665</b>	-0,152
Computer	<b>0,629</b>	-0,029
Künstlerische Tätigkeiten	<b>0,553</b>	0,136
Beschäftigung mit Kindern	0,054	<b>0,863</b>
Beschäftigung mit Familie	-0,015	<b>0,857</b>

**Tabelle 7: Lebensstil: Dimensionen der Freizeitpräferenzen – Matrix der Faktorladungen**

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Über diese Dimensionen hinaus dürfte nach früheren Analysen im Projekt StadtLeben auch die Selbstverwirklichungsorientierung von Bedeutung sein.

#### 4.4.3.3 Versorgungswege

Für Versorgungswege wird im Lebensstilbereich vor allem die familiäre Verantwortung als bedeutsam angesehen. Diese drückt sich in den erhobenen Items in den Freizeitaktivitäten ‚mit der Familie beschäftigen‘ und ‚mit den Kindern beschäftigen‘ aus. Diese korrelieren darüber hinaus positiv mit ‚Gartenarbeit‘.

Sowohl die Reliabilitätsanalyse als auch eine Faktorenanalyse führt allerdings bei zusätzlicher Berücksichtigung der Gartenarbeit zu einem ungünstigeren Ergebnis. Deshalb werden nur die Items ‚mit der Familie beschäftigen‘ und ‚mit den Kindern beschäftigen‘ berücksichtigt. Die beiden Items gehen in eine Faktorenanalyse der Freizeitpräferenzen ein und bilden gemeinsam eine Dimension ‚häuslich-familiäre Orientierung‘ (s.o.).

#### 4.4.3.4 Verkehrsmittelnutzung: MIV und ÖV

Bei der Verkehrsmittelnutzung ist zwischen MIV und ÖV einerseits sowie NMIV andererseits zu unterscheiden. Für die Nutzung des MIV und des ÖV dürften die gleichen Lebensstilindikatoren relevant sein, nur mit umgekehrten Vorzeichen. Der NMIV ist davon zu unterscheiden, weil er weniger von Verkehrssystemen und den zugehörigen subjektiven Repräsentationen (z.B. Wahrnehmungen, Einstellungen) geprägt ist als von räumlichen Strukturen und damit korrespondierenden individuellen räumlichen Orientierungen.

Unter den erhobenen Lebensstil-Items dürften für den MIV und ÖV vor allem statusorientierte Merkmale relevant sein. Deshalb wird hier der Faktor ‚Selbstverwirklichung‘ verwendet (s.o.).

#### 4.4.3.5 Verkehrsmittelnutzung: NMIV

Eine starke Nutzung des NMIV impliziert vor allem eine starke räumliche Orientierung bzw. setzt diese voraus. Diese dürfte mit einer traditionellen, ‚bescheidenen‘ Wertorientierung korrespondieren, die geringe Ansprüche an Vielfalt, Individualität und Ausgefallenheit von Konsumangeboten stellt. Dies kommt in den Lebenszielen Sicherheit und Geborgenheit, Sparsamkeit und eventuell auch ‚für andere da sein‘ (Selbstaufopferung/Altruismus als Gegensatz zur ich-bezogenen Selbstverwirklichung) zum Ausdruck. Dafür wird der Faktor ‚traditionelle Wertorientierung‘ herangezogen (siehe oben unter Berufswege).

#### 4.4.3.6 Zusammenfassung

Im Ergebnis stehen vier Lebensstilindikatoren zur Verfügung, die vier Dimensionen der Lebensstile aus dem Projekt StadtLeben stark ähneln, aber genauer auf die entsprechenden Dimensionen zugeschnitten sind und weniger von ‚unpassenden‘ Items beeinflusst werden.

- Wertorientierung ‚traditionell‘
- Wertorientierung ‚Selbstverwirklichung‘
- Freizeitpräferenzen ‚außerhäuslich‘
- Freizeitpräferenzen ‚häuslich-familiär‘

	Wert-orientierung Tradition	Wert-orientierung Selbstverwirklichung	außer-häusliche Freizeit	häuslich-familiäre Freizeit
Wertorientierung Tradition	1	0,000	-0,206	0,183
Wertorientierung Selbstverwirklichung		1	0,322	0,089
außerhäusliche Freizeit			1	0,000
häuslich-familiäre Freizeit				1

**Tabelle 8: Korrelationen zwischen den Lebensstilindikatoren**

Alle Korrelationen > 0 sind signifikant ( $\alpha = 0,01$ ).

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Zu beachten ist, dass diese Lebensstilindikatoren nicht unabhängig voneinander sind. Sie sollten deshalb nicht als ‚Dimensionen‘ oder ‚Faktoren‘ bezeichnet werden<sup>13</sup>. Die Korrelationen zwischen den Indikatoren zeigt Tabelle 8. Insbesondere zwischen der Stärke der außerhäuslichen Freizeitorientierung

<sup>13</sup> Eine gemeinsame Faktorenanalyse über alle in den beiden obigen Faktorenanalysen berücksichtigten Items wurde probeweise durchgeführt, führt aber nicht zu einer akzeptablen Trennung zwischen den Dimensionen und ist damit im Ergebnis schlecht interpretierbar.

und den beiden Wertorientierungen bestehen deutliche Korrelationen. Zwischen den beiden Wertorientierungen sowie zwischen den beiden Freizeitorientierungen jeweils untereinander bestehen keine Korrelationen, weil sie das Ergebnis gemeinsamer Faktorenanalysen sind.

#### 4.5 Lebensstil II

Wie bei der Lebenslage gilt auch beim Lebensstil, dass die Arbeit mit Faktorwerten in Strukturgleichungsmodellen unbefriedigend ist, weil sie Messfehler aus den Modellen heraus zwingt. Deshalb wurden bei der weiteren Arbeit auch bei der Abbildung des Lebensstils in den Modellen Messfehler zugelassen, d.h. der Lebensstil wurde in Form konfirmatorischer Faktorenanalysen als Messmodell in den Modellen abgebildet. Die Faktorenanalysen orientierten sich an der im vorigen Kapitel 0 beschriebenen Vorgehensweise. Alle Variablen wurden an der generellen Zustimmung- oder Ablehnungstendenz des jeweiligen Befragten normiert. Der Grund dafür ist die unterschiedlich stark ausgeprägte generelle Zustimmungsbereitschaft unter den Befragten. Dies führt dazu, dass manche Befragten alle Items abzulehnen scheinen, während andere tendenziell allen Items zustimmen. Dies kann beispielsweise bei der Gruppenbildung zu Artefakten führen ("Zustimmer" scheinen in allen Bereichen hochaktiv und positiv eingestellt zu sein). Die Normierung führt dazu, dass nur das relative Überwiegen der Zustimmung zu einem Item – gemessen an der Zustimmung der betreffenden Person zu allen anderen Items – als Zustimmung gewertet wird. Positive Werte stehen für eine überdurchschnittliche Zustimmung und negative Werte für eine überdurchschnittliche Ablehnung relativ zu allen anderen Kriterien.

Im Modell der Freizeitpräferenzen wird die auftretende schwache Korrelation zwischen außerhäuslicher Freizeit und häuslicher Freizeit ( $r=0.06$ ) nicht zugelassen. Daraus resultiert eine etwas bessere Modellanpassung, und in den späteren Strukturmodellen muss ein Parameter weniger geschätzt werden, so dass der Zwang zur Fixierung von Parametern - um die Modelle identifizierbar zu machen - an anderer Stelle gelockert wird.

#### 4.6 Wohnstandortpräferenzen

Für den theoretischen Ansatz des Projekts spielen Selbstselektionseffekte der Bevölkerung bei der Wohnstandortwahl eine zentrale Rolle. Diese werden durch die Berücksichtigung von Wohnstandortpräferenzen abgebildet. Gemeint ist damit die Wichtigkeit oder Relevanz von Eigenschaften des Standortes für die Wohnstandortentscheidungen der befragten Person.

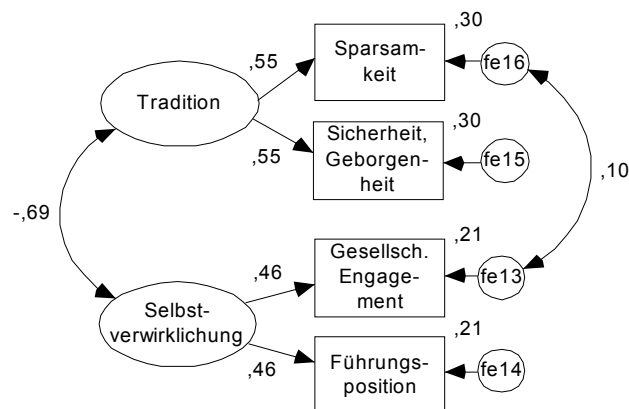


Abbildung 3: Messmodell des Lebensstils I: Wertorientierungen

CMIN/DF=0.787, RMSEA=0.000, Hoelter ( $p=0.05$ )=10193

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

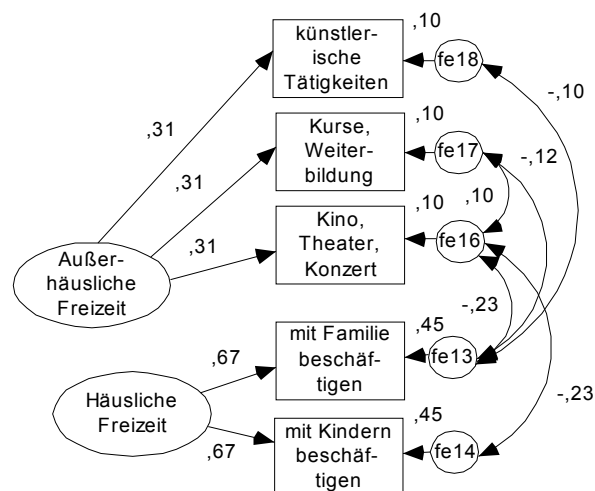


Abbildung 4: Messmodell des Lebensstils II: Freizeitpräferenzen

CMIN/DF=1.370, RMSEA=0.012, Hoelter ( $p=0.05$ )=5870

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Erfasst wurden dabei in Form fünfstufiger Skalen eine Reihe von Eigenschaften des Wohnstandorts und der Wohnung. Tabelle 9 gibt eine Übersicht der Items und ihrer Zuordnung zu Skalen, wie sie in StadtLeben vorgenommen wurde (Scheiner 2005b). Bei der Formulierung der Ausprägungen der Skalen wurde versucht, einer metrischen Skala möglichst nahe zu kommen (Rohrmann 1978). Die Bewertungen der einzelnen Eigenschaften korrelieren hochgradig miteinander.

Wie beim Lebensstil wurden die Werte an der Zustimmung- oder Ablehnungstendenz des Befragten normiert. Die Antworten nehmen dadurch einen Wertebereich von -4 bis +4 an. Dabei stehen posi-

tive Werte für eine überdurchschnittliche Bedeutung und negative Werte für eine unterdurchschnittliche Bedeutung *relativ zu allen anderen Kriterien.*

Thema	Items: Subjektive Wichtigkeit von...
Erreichbarkeit des Zentrums	Erreichbarkeit des Stadtzentrums, Erreichbarkeit mit Bussen und Bahnen, Erreichbarkeit des Arbeits-/Ausbildungsplatzes
Versorgungsinfrastruktur	Einkaufsmöglichkeiten, Dienstleistungen, Freizeitmöglichkeiten für Erwachsene
Kinder	Freizeitmöglichkeiten für Jugendliche, Spiel- und Freizeitmöglichkeiten für Kinder, Kindergärten/Schulen
Wohnumfeld	Nachbarschaft, Sicherheit vor Kriminalität/Vandalismus, Störung durch Lärm/ Abgase/Schmutz, Zustand der Parks und Grünflächen
Wohnung	Ausstattung, baulicher Zustand, Größe und Schnitt, Höhe der Kosten.
Pkw-Erreichbarkeit	Anbindung an Autobahnen / Schnellstraßen, Parkplätze / Garagen

**Tabelle 9: Items der Wohnstandortpräferenzen**

Quelle:: Projekt StadtLeben.

Aus Gründen der Sparsamkeit werden in einem Modell nur ausgewählte Aspekte der Wohnstandortpräferenzen verwendet. Die Auswahl orientiert sich an theoretischen Überlegungen. Für Berufswege wird die Wichtigkeit der Erreichbarkeit des Arbeitsplatzes verwendet, für Versorgungswege findet eine latente Variable 'Bedeutung von Nähe zu Versorgung' Verwendung, die auf den Items 'Wichtigkeit von Nähe zu Einkaufsmöglichkeiten' und 'Dienstleistungen' basiert. Für Freizeitwege wird die Wichtigkeit der Ausstattung des Wohngebietes mit Freizeitgelegenheiten für Erwachsene eingesetzt, für die Nutzung des MIV und des ÖV das ÖPNV-Angebot. Für die NMIV-Nutzung wird nicht die ÖPNV-Angebot verwendet, sondern die Nähe zu Einkaufsmöglichkeiten und Dienstleistungen, weil für Fußwege weniger das Verkehrssystem als die nahräumliche Ausstattung des Wohngebietes von Bedeutung ist.

## 4.7 Objektive Raumstruktur

Neben den Wohnstandortpräferenzen der Befragten müssen auch objektive raumstrukturelle Merkmale des Wohnstandorts in den Analysen berücksichtigt werden. Ansonsten bestünde die Gefahr, dass existierende Effekte der Raumstruktur bzw. der tatsächlich getroffenen Standortentscheidung auf das Verkehrshandeln fälschlicherweise als Effekte der Standortpräferenzen – also der Selbstselektion der Bevölkerung – interpretiert würden.

Im Folgenden wird die Aufbereitung der vorhandenen Daten aus den Vor-Ort-Erhebungen und aus weiteren Quellen beschrieben. Die Indikatoren umfassen folgende Bereiche: Arbeitsplätze, Schulen,

Versorgungseinrichtungen, Freizeitangebote, ÖPNV-Angebote und Lage in der Region.

### 4.7.1 Arbeitsplatzangebot

#### 4.7.1.1 Konstruktion eines Indikators Arbeitsplatzerreichbarkeit

Das Arbeitsplatzangebot wird, wie in Kapitel 3.2.1 beschrieben, über die räumliche Verteilung der Arbeitsplätze beschrieben.

Aus den vorhandenen Daten wurde mit Hilfe eines Gravitationsansatzes ein Indikator der Arbeitsplatzerreichbarkeit konstruiert. Dabei wird die Zahl der Arbeitsplätze in einer Raumeinheit (Gemeinde, Ortsteil oder Verkehrszelle) als Gelegenheitenpotenzial betrachtet, das umgekehrt zur Entfernung vom jeweiligen Untersuchungsgebiet gewichtet wird. Als Distanzmaß wird die kilometrische Entfernung verwendet. Ein adäquateres Maß wäre die Zeitdistanz, die für Pendler eine wichtigere Rolle spielt; eine Matrix der Zeitdistanzen für die Region Köln steht jedoch nicht zur Verfügung.

#### 4.7.1.2 Bestimmung eines Koeffizienten für den Raumwiderstand

Die Gewichtung der erreichbaren Arbeitsplätze je nach Entfernung hängt vom Raumwiderstand ab. Der Verlauf einer entsprechenden Gewichtungsfunktion lässt sich durch Einführung eines Koeffizienten beeinflussen, der den Widerstand der Distanz spezifiziert.

Für die Stärke des Koeffizienten gibt es keine allgemeinen Richtwerte. Als klassisch gilt  $\alpha=2$ , d.h. die Entfernung wird quadriert. Dafür gibt es allerdings keinen theoretischen Grund, so dass der Koeffizient möglichst von Fall zu Fall empirisch ermittelt werden sollte (Titheridge 2000). Bram und McKay (2005) ermitteln für New York City gar Koeffizienten von  $\alpha=3,25$  (1980) und  $\alpha=3,10$  (2000). Dagegen errechnet Johnson (2003) für Atlanta, Boston und Los Angeles Werte zwischen 0,09 und 0,15 für Beschäftigte mit niedrigerem Einkommen und zwischen 0,05 und 0,07 für Beschäftigte mit höherem Einkommen. Die niedrigeren Werte der zweiten Gruppe spiegeln deren geringere Distanzempfindlichkeit wider.

Im vorliegenden Fall wurde näherungsweise ein geeigneter Koeffizient empirisch ermittelt. Dieser lässt sich bestimmen, indem die Auswahlwahrscheinlichkeit eines Arbeitsplatzes durch eine erwerbstätige Person in Abhängigkeit von der Entfernung vom Wohnstandort bestimmt wird. Diese Auswahlwahrscheinlichkeit sollte mit zunehmender Entfernung geringer werden; zu beantworten ist die Frage: in welchem Maß geringer?

Aufgrund des hohen Anteils innergemeindlicher Wege lässt sich eine hinreichend gute Annäherung

nur für Gebiete ermitteln, für die eine kleinräumliche (innergemeindliche) Differenzierung des Arbeitsplatzangebots in einem Radius von wenigstens 10 km um das Quellgebiet der Wege zur Verfügung steht. Dies ist im Rahmen der StadtLeben-Daten für Ehrenfeld und Nippes der Fall. Eine Beschränkung auf diese beiden Gebieten ist allerdings mit zwei Problemen behaftet: erstens die kleinen Teilstichproben, zweitens der sehr spezifische Charakter der gründerzeitlichen Gebiete, die vor allem für Existenzgründer und andere junge, vorwiegend akademische Erwerbstätige attraktiv sind, was zu einer ebenfalls spezifischen Verteilung der Arbeitsplatzstandorte führt.

Deshalb wurden für die Ermittlung des Koeffizienten für den Distanzwiderstand sowohl Analysen für Ehrenfeld und Nippes als auch für alle Untersuchungsgebiete gemeinsam durchgeführt.

Zunächst wird dabei die Wegelängenverteilung der Arbeitswege bestimmt. Zum anderen wird die Verteilung der Arbeitsplätze in Köln und der Region auf Entfernungsklassen von jedem Untersuchungsgebiet aus bestimmt. Aus beiden Werten lässt sich für jede Entfernungsklasse das Verhältnis zwischen dem Anteil der Wege in der jeweiligen Klasse an allen Arbeitswegen und dem Anteil der Arbeitsplätze in der gleichen Klasse an allen Arbeitsplätzen bestimmen. Dies spiegelt das Verhältnis zwischen Arbeitsplatznachfrage und Arbeitsplatzangebot in Abhängigkeit von der Entfernung vom Wohnstandort wider.

a) Ergebnis für Ehrenfeld und Nippes

Obwohl es sich aufgrund der kleinen Teilstichprobe um eine recht grobe Schätzung handelt, ergibt sich im Wesentlichen ein Funktionsverlauf, der den Erwartungen entspricht (Abbildung 5). Lediglich bei etwa zwei bis vier Kilometer Entfernung zeigt sich in der Auswahlwahrscheinlichkeit für einen Arbeitsplatz ein "Einbruch". Dies entspricht der Entfernung zur Kölner Altstadt mit ihrer äußerst hohen Arbeitsplatzdichte. Daraus ist nicht zu schließen, dass nur wenige Ehrenfelder bzw. Nippeser in der Altstadt arbeiten. Vielmehr dürfte es so zu verstehen sein, dass die Altstadt mit ihrer extrem hohen Arbeitsplatzzentralität für die gesamte Region Köln ein äußerst attraktives Potenzial darstellt, so dass die Wahrscheinlichkeit für die Auswahl eines Arbeitsplatzes in der Altstadt durch einen Ehrenfelder bzw. Nippeser eher gering ist.

Durch Variieren des Raumwiderstandskoeffizienten lässt sich nun eine Anpassung der Raumwiderstandsfunktion an den empirischen Verlauf vornehmen. Die insgesamt beste Anpassung liefert der Koeffizient  $\alpha=1$  (Abbildung 5).

Nach diesem recht niedrigen Koeffizienten scheint die Distanzempfindlichkeit geringer als zunächst erwartet. Dies ist auf die hier verwendete kleinräumliche Auflösung zurückzuführen. Die meisten Pendlerstudien, die auf Gravitationsmodellen beruhen, verwenden eine deutlich grobkörnigere räumliche

Auflösung (etwa Gemeinden in der BRD, Counties in den USA). Da dann bereits die nächst benachbarten Raumeinheiten für relativ große Entfernungen stehen, nimmt die Auswahlwahrscheinlichkeit für einen Arbeitsplatz bereits bei 'geringer' Entfernung (nämlich in den benachbarten Raumeinheiten) stark ab. Dies ist hier – bei einer wesentlich feineren räumlichen Auflösung – nicht im gleichen Maße der Fall. Insofern ist der vergleichsweise niedrige Koeffizient keine Überraschung.

b) Ergebnis für alle Untersuchungsgebiete

Werden alle Untersuchungsgebiete gemeinsam betrachtet, so ergibt sich eine bessere Kurvenanpassung, wenn ein höherer Koeffizient gewählt wird (Abbildung 6). Im Entfernungsbereich bis zu vier Kilometern liefert der Koeffizient  $\alpha=2$  die beste Anpassung, im weiteren Radius schätzt dieser allerdings zu niedrige Werte, so dass  $\alpha=1,5$  angemessener erscheint.

Wie oben bereits erwähnt, ist dies nicht unbedingt inhaltlich zu interpretieren. Ehrenfeld und Nippes sind die einzigen Untersuchungsgebiete, für die eine kleinräumliche Differenzierung des Arbeitsplatzangebots in einem recht großen Radius vorliegt. Die eher grobe Differenzierung auf Gemeindeebene im Kölner Umland führt bei der gemeinsamen Analyse für alle Gebiete zu einer vermeintlich höheren Distanzempfindlichkeit und damit zu einem höheren Koeffizienten.

Bemerkenswert ist über diese Ergebnisse hinaus der Vergleich der Endresultate der Analysen: die Werte der Arbeitsplatzerreichbarkeit für die Untersuchungsgebiete bei verschiedenen Raumwiderständen (Tabelle 10).

	$\alpha=2$	$\alpha=1,5$	$\alpha=1$
Ehrenfeld	-0,22	0,40	1,59
Nippes	-0,16	0,38	1,45
Stammheim	-0,51	-0,44	0,29
Longerich	-0,50	-0,45	0,19
Esch	-0,56	-0,68	-0,37
Zündorf	-0,57	-0,73	-0,70
Kerpen	0,48	0,21	-0,78
Sindorf	2,69	2,53	0,61
Overath	-0,08	-0,38	-1,07
Heiligenhaus	-0,56	-0,85	-1,22

**Tabelle 10: Z-Standardisierte Werte der Arbeitsplatzerreichbarkeit in den Untersuchungsgebieten für drei verschiedene Koeffizienten des Raumwiderstands**

Quelle: eigene Analysen.

Bei einem Raumwiderstandskoeffizienten von  $\alpha=2$  ragt das Gebiet Sindorf mit einer extrem hohen Arbeitsplatzerreichbarkeit gegenüber allen anderen Gebieten heraus. Dies ist Folge des Gewerbegebiets Sindorf, in dem rund zwei Drittel der Arbeitsplätze der Stadt Kerpen liegen. Weil mit dem höheren Ko-

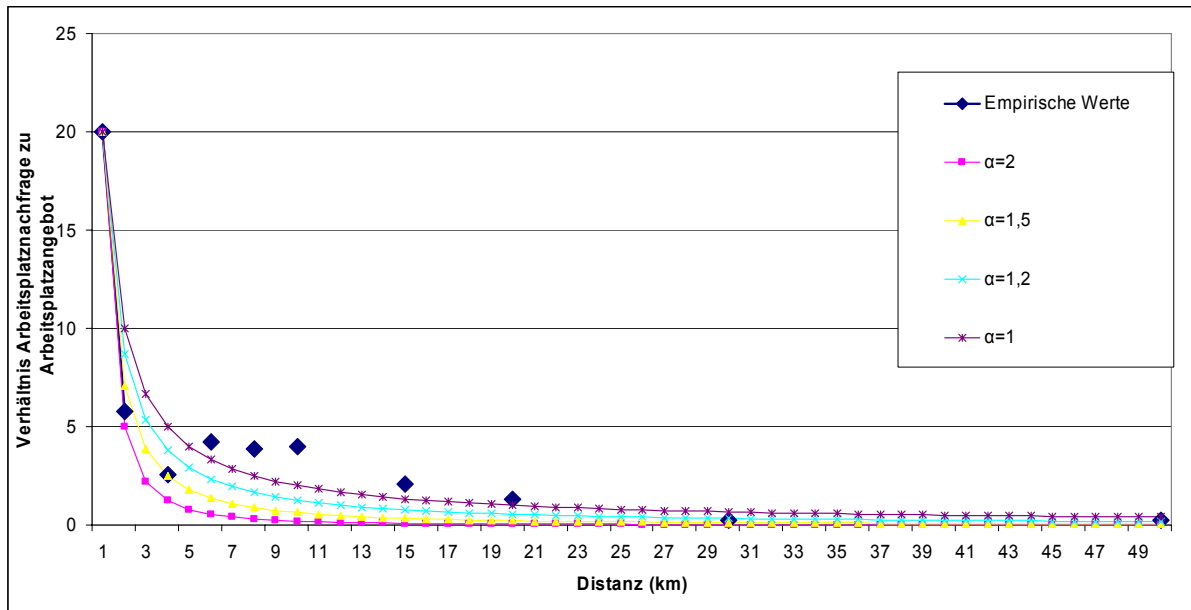


Abbildung 5: Verlauf der Distanzwiderstandsfunktion (Ehrenfeld und Nippes): Empirische Werte und Kurvenanpassung mit verschiedenen Widerstandskoeffizienten

Quelle: eigene Analysen.

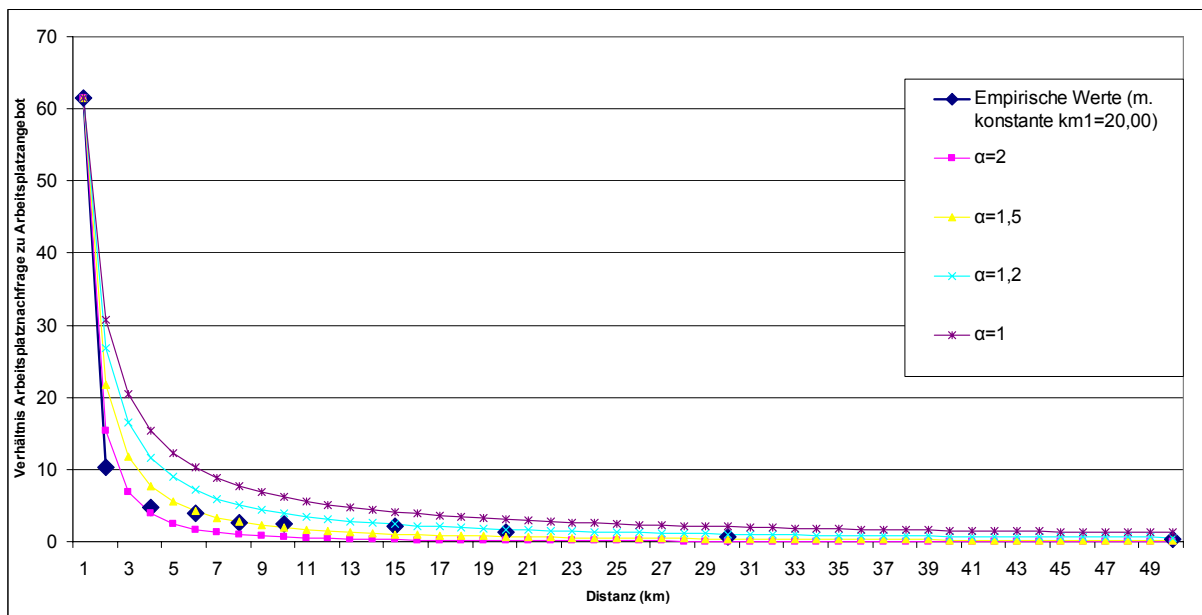


Abbildung 6: Verlauf der Distanzwiderstandsfunktion (alle Untersuchungsgebiete): Empirische Werte und Kurvenanpassung mit verschiedenen Widerstandskoeffizienten

Quelle: eigene Analysen.

effizienten  $\alpha=2$  eine sehr hohe Distanzempfindlichkeit angenommen wird, erhalten diese in Sindorf gelegenen Arbeitsplätze gegenüber den etwas weiter entfernten Arbeitsplätzen ein sehr hohes Gewicht. Dies schlägt sich auch noch im nahe gelegenen Kerpen nieder, das (weit hinter Sindorf) den zweitbesten Wert aufweist. Beim niedrigeren Koeffizienten  $\alpha=1,5$  wird dieser Ausreißer nur geringfügig abge-

mildert. Erst bei  $\alpha=1$  sind die Verhältnisse ausgewogener und entsprechen eher Plausibilitätserwartungen. Herausragend sind nun Ehrenfeld und Nippes als innerstädtische Gebiete der Kernstadt Köln. Dennoch behält Sindorf den höchsten Wert aller suburbaner Untersuchungsgebiete und liegt in der Arbeitsplatzreichbarkeit besser als die innenstadtnahen Gebiete Stammheim und Longerich in Köln.

Dieses Ergebnis scheint plausibel, so dass der Gravitationsansatz mit dem Koeffizienten  $\alpha=1$  verwendet wird.

#### 4.7.2 Schulen

Aus den Erhebungen liegen Adressen der relevanten Schulen vor. Diesen wurden geographische Koordinaten zugeordnet, aus denen die Distanzen der Wohnstandorte der Befragten zur nächsten Hauptschule, zur nächsten Realschule und zum nächsten Gymnasium ermittelt.

Im Ergebnis treten je nach Schulart und Gebiet teilweise erhebliche Entfernungen auf. Selbst in den innenstadtnahen Gebieten Stammheim und Longenrich beträgt die Luftlinienentfernung zur nächsten Realschule rund 3 km, in Longenrich ist auch das nächste Gymnasium über 3 km entfernt.

#### 4.7.3 Versorgungs- und Freizeitgelegenheiten

Wie erwähnt, liegen für Versorgungsangebote (Einzelhandel, Dienstleistungen) und Freizeitangebote umfangreiche Daten aus eigenen Erhebungen vor. Darüber hinaus liegt ein ordinalskaliertes Indikator der Erholungsqualität der Untersuchungsgebiete vor.

Aus den Daten wurden eine Reihe von Indikatoren gebildet, die im Folgenden genannt werden. Die auf einer Anzahl basierenden Indikatoren beziehen sich jeweils auf einen Radius von 650 m Luftlinie um den Wohnstandort der befragten Person, der aus geographischen Koordinaten ermittelt wurde.

- Distanz zum nächstgelegenen größeren Lebensmittelgeschäft (Supermarkt, Verbrauchermarkt)
- Anzahl größerer Lebensmittelgeschäfte (Supermarkt, Verbrauchermarkt)
- Anzahl weiterer Lebensmittelgeschäfte
- Anzahl weiterer Geschäfte
- Anzahl Dienstleistungsbetriebe für den täglichen und periodischen Bedarf (Bank, Post, Schneider, Friseur, Kopiergeschäft, Reisebüro usw.; ohne medizinische Versorgung, Handwerk und unternehmensbezogene Dienstleistungen) inkl. öffentliche Verwaltung<sup>14</sup>
- Anzahl medizinische Dienstleistungen (Ärzte, Krankenhäuser, Heilpraktiker, Pflegedienste usw.)
- Anzahl Freizeitgelegenheiten unter Einschluss informeller Treffpunkte

- Anzahl gastronomische Einrichtungen (Restaurants, Gaststätten, Cafés usw.).

Aufgrund der großen Bedeutung privater Besuche für die außerhäusliche Freizeit wird ähnlich der Erreichbarkeit von Arbeitsplätzen ergänzend auch die Erreichbarkeit von Bevölkerungspotenzialen mit einem Gravitationsansatz für jedes Quartier ermittelt. Dies entspricht der Vorgehensweise in Verkehrsberechnungsmodellen für den Freizeitverkehr. Diesem Vorgehen liegt der Gedanke zugrunde, dass die Bevölkerung selbst ein Gelegenheitenpotenzial in der Freizeit darstellt. Die räumliche Verteilung der Bevölkerung erfolgt innerhalb Kölns anhand der Stadtteile; in der unmittelbaren Umgebung der Untersuchungsgebiete wird noch genauer nach Stadtvierteln differenziert. Außerhalb Kölns wird nach Gemeinden differenziert, in Kerpen und Overath zusätzlich nach Orts- bzw. Stadtteilen. Als Widerstandskoeffizient wird wiederum  $\alpha=1$  gewählt, weil höhere Widerstände zu unplausiblen Ergebnissen führen (Tabelle 11). Insbesondere führen höhere Widerstände zu extrem hohen Werten für Kerpen und Sindorf, weil dort innerhalb der Ortsteile keine kleinräumliche Verteilung der Bevölkerung vorliegt und somit die Gesamtbevölkerung des jeweiligen Ortsteils in extrem geringer Entfernung erreichbar scheint, während in Köln durch die genauere Differenzierung selbst in den hochverdichteten Quartieren Ehrenfeld und Nippes die erreichbare Bevölkerung eher gering ausfällt.

	$\alpha=2$	$\alpha=1,5$	$\alpha=1$
Ehrenfeld	-0,48	-0,03	1,45
Nippes	-0,61	-0,20	1,46
Stammheim	-0,74	-0,67	0,55
Longenrich	-0,68	-0,66	0,31
Esch	-0,45	-0,54	-0,11
Zündorf	-0,78	-1,01	-1,00
Kerpen	1,70	1,68	-0,21
Sindorf	1,82	1,84	0,09
Overath	0,53	0,28	-1,05
Heiligenhaus	-0,32	-0,69	-1,48

**Tabelle 11: Z-Standardisierte Werte der Erreichbarkeit von Bevölkerungspotenzialen in den Untersuchungsgebieten für drei verschiedene Koeffizienten des Raumwiderstands**

Quelle: eigene Analysen.

#### 4.7.4 ÖPNV-Angebot

Zur Abbildung des ÖPNV-Angebots liegt ein Indikator „ÖPNV-Systemqualität“ vor, der von der RWTH Aachen, Institut für Stadtbauwesen und Stadtverkehr, erstellt wurde. Dabei handelt es sich um eine ordinalskalierte Variable, die auf den unterschiedlichen Funktionen der ÖPNV-Verkehrsträger beruht:

- Der Bus bzw. (im Umland) das Anrufsammeltaxi (AST) nimmt vorwiegend kleinräumliche Erschließungsfunktionen wahr,

<sup>14</sup> Diese umfasst die Bezirksämter Ehrenfeld und Nippes sowie je zwei Standorte der Stadtverwaltungen Kerpen und Overath. In den anderen Untersuchungsgebieten sind von der öffentlichen Verwaltung allenfalls Polizei oder Feuerwehr im kartierten Radius enthalten. Beides wurde nicht berücksichtigt.

- Die S-Bahn und Regionalbahn nimmt vorwiegend die regionale (großräumliche) Verbindungsfunktion wahr,
- Die U-Bahn liegt dazwischen und dient der mesoräumlichen Erschließung und Verbindung.

Zur Bildung der Variable wurde das ÖPNV-Angebot in den Gebieten in fünf Stufen klassifiziert:

- 1 = keine ÖPNV-Bedienung (kommt in den untersuchten Gebieten nicht vor)
- 2 = ÖPNV-Bedienung mit einer Art von Verkehrsträger (d.h. entweder Bus/AST oder S-Bahn/Bahn oder U-Bahn) (Esch, Kerpen-Stadt, Heiligenhaus; dies sind die nur durch den Bus erschlossenen Untersuchungsgebiete)
- 3 = Bedienung z.B. mit U-Bahn und S-Bahn, d.h. ohne kleinräumige Erschließungsfunktion oder mit U-Bahn und Bus, d.h. ohne großräumige Verbindungsfunktion (Zündorf; dort fehlt die Buserschließung)
- 4 = Bedienung mit Bus und U- oder S-Bahn, d.h. sowohl Erschließung als auch klein- oder großräumige Verbindung (Stammheim, Sindorf, Overath-Stadt)
- 5 = Bedienung mit allen drei Arten von Verkehrsträgern, d.h. umfassendes integriertes Angebot (Ehrenfeld, Nippes, Longerich).

Darüber hinaus liegen für jede befragte Person auf der Individualebene folgende Indikatoren vor: Distanz zur nächsten Fern- oder S-Bahn-Haltestelle; Distanz zur nächsten U-Bahn-Haltestelle; Distanz zur nächsten Bushaltestelle; Distanz zur nächsten AST-Haltestelle.

Aufgrund der ungleichen Verteilung der verschiedenen Systeme (AST nur im Umland, U-Bahn nur in Köln) werden diese Merkmale zusammengefasst zu zwei Variablen:

- Distanz zum nächsten Schienenhaltepunkt
- Distanz zur nächsten ÖPNV-Haltestelle (ohne Unterscheidung zwischen Straße und Schiene).

#### 4.7.5 Lage der Gebiete

Über die Ausstattung der Untersuchungsgebiete und ihres Umfeldes hinaus liegen aus Erhebungen des Projekts StadtLeben mehrere Indikatoren der relativen Lage der Gebiete vor, nämlich die Fahrzeit vom Mittelpunkt der Gebiete ins Zentrum Kölns (Neumarkt; differenziert nach MIV und ÖV), die Entfernung zum Zentrum Kölns sowie die Entfernung zum nächsten Mittelzentrum bzw. Bezirkszentrum Kölns.

#### 4.7.6 Bildung raumstruktureller Indikatoren

Die folgende Tabelle gibt einen Überblick über die vorliegenden Indikatoren zur Beschreibung der Raumstruktur. Diese sind nun auf inhaltlich sinnvolle

Weise zu verdichten. Ziel ist eine möglichst geringe Zahl von stetigen Indikatoren, die allerdings so differenziert sein sollen, dass sie separate Analysen für Berufs-, Ausbildungs-, Versorgungs- und Freizeitwege sowie für die Verkehrsmittelnutzung sinnvoll möglich machen. Detaillierte Indikatoren zur Untersuchung besonderer Wegezwecke (z.B. Gastronomiebesuche) werden zusätzlich verwendet.

Die Korrelationen zwischen den Variablen sind häufig sehr hoch. Eine Faktorenanalyse unter Einschluss der meisten Variablen führt zu einer hohen Varianzaufklärung von 90% und der Extraktion von fünf Faktoren. Diese sind jedoch nicht in allen Fällen inhaltlich sinnvoll interpretierbar. Ziel ist es zudem nicht, die Raumstruktur in Form unabhängiger Dimensionen (Faktoren) zu beschreiben. Vielmehr geht es darum, für verschiedene Aspekte der Verkehrsnachfrage auf jeweils adäquate Indikatoren der Raumstruktur zurückgreifen zu können. Insofern scheint eher eine inhaltlich begründete Verdichtung der Variablen sinnvoll. Dafür werden die Variablen standardisiert und zu Skalen zusammengefasst. Die statistische Sinnfälligkeit der Skalen wird mit Cronbach's  $\alpha$  sowie mit Faktorenanalysen über die *in die jeweilige Skala eingehenden* Variablen geprüft. Die Faktoren, die dabei extrahiert werden, sind also nicht unabhängig voneinander. Als sinnvoll gilt eine Skala dann, wenn sich die darin eingehenden Variablen mit hinreichender Varianzaufklärung zu *einem* Faktor zusammenfassen lassen und wenn Cronbach's  $\alpha$  einen akzeptablen Wert von mindestens  $\alpha=0,60$  erreicht. Die Variablen sind dafür vorher zu standardisieren. Aufgrund der starken Korrelationen und der entsprechenden hohen Reliabilität der Indikatoren erscheint es gerechtfertigt, hier und bei den anderen raumstrukturellen Variablen von Messfehlerfreiheit auszugehen und keine latenten Variablen zu bilden.

**ÖPNV:** Eine Analyse auf der Grundlage von Entfernung zum nächsten Schienenhaltepunkt, Entfernung zum nächsten ÖPNV-Haltepunkt und ÖPNV-Systemqualität ergibt ein inhaltlich plausibles Ergebnis. Allerdings wird die Güte des Indikators durch die Unspezifität der Entfernung zum nächsten ÖPNV-Haltepunkt beeinträchtigt, die sich aus der mehr oder weniger flächendeckenden Erschließung mit dem Bus bzw. AST ergibt. Deshalb wird diese Variable aus der Analyse entfernt, so dass nur zwei standardisierte Variablen in die Skala eingehen: Entfernung zum nächsten Schienenhaltepunkt und ÖPNV-Systemqualität. Dies scheint eine hinreichend genaue Abbildung des ÖPNV-Angebots zu gewährleisten und beinhaltet sowohl die Entfernung zum Haltepunkt als auch die Angebotsvielfalt in Form einer mehr oder weniger großen Zahl von Teilsystemen mit unterschiedlichen Funktionen (Cronbach's  $\alpha=0,93$ , Varianzaufklärung 93,4%).

**Arbeitsplatzerreichbarkeit:** bleibt wie oben beschrieben.

Variablen- name	Inhalt	Herkunft der Variable	Skalierungs niveau	Bezug
	ÖPNV			
d_sbahn	Distanz nächste Fern-/S-Bahn	EB	Metrisch	Person
d_ubahn	Distanz nächste U-Bahn	EB	Metrisch	Person
d_bus	Distanz nächste Bushaltestelle	EB	Metrisch	Person
d_ast	Distanz nächste AST-Haltestelle	EB	Metrisch	Person
D_bahn	Distanz nächster Schienenhaltepunkt	EB	Metrisch	Person
D_öv	Distanz nächste ÖPNV-Haltestelle	EB	Metrisch	Person
ÖPNV_2	ÖPNV-Systemqualität	SL	Ordinal	UG
öpnv2_di	ÖPNV-Systemqualität	SL	Dichotom	UG
	Arbeitsplätze			
Aplätze	Anzahl Arbeitsplätze im UG	SL	Metrisch	UG
Aplatzdi	Arbeitsplatzdichte im UG (Arbeitsplätze/qkm)	SL	Metrisch	UG
Apl1	Arbeitsplatzerreichbarkeit	EB/PTV	Metrisch	UG
	Schulen			
dmn_hs	Distanz nächste Hauptschule	EB	Metrisch	Person
dmn_rs	Distanz nächste Realschule	EB	Metrisch	Person
dmn_gy	Distanz nächstes Gymnasium	EB	Metrisch	Person
	Einzelhandel, Dienstleistungen			
dmn_sup	Distanz nächster Super-/Verbrauchermarkt	EB	Metrisch	Person
z_sup650	Anzahl Super-/Verbrauchermärkte im Radius von 650m Luftlinie	EB	Metrisch	Person
z_leb650	Anzahl Lebensmittelgeschäfte (ohne Supermärkte) Radius 650m Luftlinie	EB	Metrisch	Person
z_eh650	Anzahl Geschäfte (ohne Lebensmittel) Radius 650m Luftlinie	EB	Metrisch	Person
z_dl650	Anzahl Dienstleistungen (ohne Medizin) Radius 650m Luftlinie	EB	Metrisch	Person
z_med650	Anzahl medizinische Dienstleistungen Radius 650m Luftlinie	EB	Metrisch	Person
vers_di	Versorgungsstruktur	SL	Dichotom	UG
Bs_800	Behaviour settings im 800m-Radius	SL	Metrisch	UG*
Bs_800ek	Behaviour settings Einkauf im 800m-Radius	SL	Metrisch	UG*
Bs_dicht	Behaviour setting Dichte im UG (BS/ha)	SL		
Fac_Zen	Faktor Zentralität	EB	Metrisch	UG
Fac_Dimi	Faktor Dichte und Mischung	EB	Metrisch	UG
	Freizeit			
z_frz650	Anzahl Freizeiteinrichtungen (ohne Grünflächen, Gastronomie) Radius 650m Luftlinie	EB	metrisch	Person
z_gas650	Anzahl Gastronomische Einrichtungen Radius 650m Luftlinie	EB	Metrisch	Person
grün	Erholungsqualität (Grünflächen)	Experten- gespräch	Ordinal	UG
ew1	Erreichbarkeit von Bevölkerungspotenzial	EB	metrisch	UG
dicht_m	Bevölkerungsdichte	SL	Metrisch	UG
dicht_di	Bevölkerungsdichte	SL	Dichotom	UG
freiz_di	Freizeitstruktur	SL	Dichotom	UG
Bs_800fz	Behaviour settings Freizeit im 800m-Radius	SL	metrisch	UG*
	Lage			
mivkö_m	Pkw-Fahrzeit nach Köln-Neumarkt	SL	Metrisch	UG
mivkö_di	Pkw-Fahrzeit nach Köln-Neumarkt	SL	Dichotom	UG
övkö_m	ÖV-Fahrzeit nach Köln-Neumarkt	SL	Metrisch	UG
övkö_di	ÖV-Fahrzeit nach Köln-Neumarkt	SL	Dichotom	UG
nzent_m	Distanz zum nächsten Zentrum (km)	SL	Metrisch	UG
nzent_di	Distanz zum nächsten Zentrum (km)	SL	dichotom	UG
km_kö_m	Distanz UG-Zentrum bis Köln-Zentrum	SL	Metrisch	UG
	Sonstiges			
motgr_m	Motorisierungsgrad	SL	Metrisch	UG
motgr_di	Motorisierungsgrad	SL	dichotom	UG

**Tabelle 12: Übersicht über raumstrukturelle Variablen**

Abkürzungen: EB: Eigene Berechnungen, SL: StadtLeben, UG: Bezogen auf das Untersuchungsgebiet; Pers: Personenbezogen; PTV: Firma PTV, Karlsruhe (s. Kapitel 3.2.1)

\* Mittelwert aus ursprünglich personenbezogenen Daten



**Schulen:** Die Entfernungen zu den drei Schultypen Hauptschule, Realschule und Gymnasium korrelieren nur teilweise miteinander. Sie korrelieren auch nur mäßig mit der Ausstattung an anderen zentralörtlichen Einrichtungen (Einzelhandel, Dienstleistungen, Werte etwa zwischen  $r=-0,3$  und  $r=-0,6$ ) und der Arbeitsplatzerreichbarkeit (Werte zwischen  $r=-0,12$  und  $r=-0,44$ ). Da in der Stichprobe nur Personen ab 16 Jahre sind, wird für entsprechende Analysen der Schüler der Wert für die Distanz zum nächsten Gymnasium herangezogen<sup>15</sup>.

**Versorgungseinrichtungen (Einzelhandel, Dienstleistungen):** Die Werte für die verschiedenen Kategorien von Geschäften und Dienstleistungsangeboten korrelieren stark miteinander und werden deshalb standardisiert und zu einer Skala zusammengefasst. Dafür wird jeweils die Anzahl an Einrichtungen im 650m-Radius um die Wohnung herangezogen, differenziert nach Super-/Verbrauchermärkten, sonstigem Lebensmitteleinzelhandel, sonstigem Einzelhandel, medizinischen Dienstleistungen, sonstigen Dienstleistungen. Die Entfernung zum nächstgelegenen Super-/Verbrauchermarkt wird ausgeklammert, weil sie die Güte der Skala deutlich beeinträchtigt. Das Ergebnis zeigt, dass die Zusammenfassung sinnvoll ist. Eine Faktorenanalyse führt zu einem Faktor mit einer Varianzaufklärung von 91,8%, Cronbach's  $\alpha=0,98$ .

**Freizeitangebote:** Analog wird bei den Freizeitangeboten verfahren. In die Analyse eingeschlossen werden Freizeiteinrichtungen (Gastronomie und sonstige), die Erreichbarkeit von Bevölkerungspotenzial sowie die Qualität von Erholungsflächen (stark negative Korrelation zu den anderen Werten, wird daher umgedreht). Eine Faktorenanalyse führt wiederum zu einem Faktor mit einer Varianzaufklärung von 77,6%. Der Reliabilitätstest führt zu einem Wert von Cronbach's  $\alpha=0,90$ .

**Lage in der Region:** Schließlich wird ein Lage-Indikator gebildet, der auf der Fahrzeit in das Kölner Zentrum (separat für MIV und ÖPNV, aus Routenplaner und Fahrplan) sowie der Kilometer-Entfernung zum Kölner Zentrum basiert. Die drei Werte sind stark korreliert, so dass eine Faktorenanalyse zu einem Faktor mit einer Varianzaufklärung von 92,9% führt, Cronbach's  $\alpha=0,96$ . Diese Variable drückt in umgekehrter Weise die Erreichbarkeit Kölns aus (positive Werte für periphere Gebiete).

Die Ergebnisse für Freizeit- und Versorgungsangebote, Lage und ÖPNV-Angebot sind in Tabelle 13 dargestellt. Aufgrund der Standardisierung der Ausgangsvariablen ähneln sie in der Skalierung Faktoren (Mittelwert=0, Standardabweichung=1). Sie

sind allerdings aufgrund des inhaltlich begründeten ‚Inputs‘ gegenüber Faktoren leichter interpretierbar. Die Ergebnisse verdeutlichen, dass die Bildung von Skalen aus mehreren Variablen zum Teil zu stärker differenzierten Ergebnissen führt als die Arbeit mit den Originalvariablen. So unterscheiden sich beispielsweise Kerpen Stadt und Sindorf in der tatsächlichen Entfernung zu Köln praktisch nicht, während bei der Skala „Entfernung nach Köln“ Sindorf aufgrund der geringeren Reisezeit mit dem ÖPNV und dem Pkw (Nähe zur Autobahn, Schienenanschluss) deutlich besser abschneidet.

Untersuchungsgebiet	Freizeitangebot	Einzelhandel + Dienstleistungen	Lage	ÖPNV-Qualität
Ehrenfeld	1,5204	2,1844	-1,6567	1,0310
Nippes	1,7161	1,5401	-1,5556	,9893
Stammheim	,1735	-,4640	-,7123	,3616
Longerich	-,1327	-,3934	-,6956	1,0021
Esch	-,3700	-,9171	-,0909	-1,2652
Zündorf	-,5881	-,4387	,1093	,0080
Kerpen Stadt	,0443	,0163	1,0088	-1,6046
Sindorf	-,0234	-,4031	,7424	,3190
Overath Stadt	-,8241	-,0180	,9242	,4608
Heiligenhaus	-,9794	-,7981	,9312	-1,2424
Insgesamt	,0000	,0000	,0000	,0000

**Tabelle 13: Raumstrukturelle Indikatoren (Mittelwerte für die Gebiete)**

Quelle: eigene Analysen.

## 4.8 Wohnmobilität und Standortwahl

Die getroffene Wohnstandortentscheidung eines Haushalts drückt sich in raumstrukturellen Merkmalen des gewählten Wohnstandorts aus. Insofern lässt sich die Wohnstandortwahl durch raumstrukturelle Merkmale des Standorts beschreiben. Die Grundlagen der Entscheidung dagegen bilden Standortpräferenzen, die in Kapitel 0 beschrieben wurden.

Allerdings liegen faktisch getroffene Standortentscheidungen zeitlich oft lange zurück, sind persistent und bleiben aufgrund der hohen Transaktionskosten einer Standortverlagerung auch bei einer Änderung der Standortpräferenzen bestehen.

Für ergänzende Analysen werden deshalb Variablen erstellt, die den Typ einer Standortverlagerung beschreiben (‚Verflechtungstyp‘) und sowohl den gegenwärtigen als auch den vorherigen oder einen früheren Wohnstandort berücksichtigen. Zudem zeigen sich in Verkehrsuntersuchungen immer wieder individuelle biographische Bindungen an frühere Wohnstandorte, die Auswirkungen auf die räumliche Struktur der Verkehrsnachfrage und auf die zurückgelegten Distanzen besitzen. So lässt sich anhand der StadtLeben-Daten zeigen, dass die langen Dis-

<sup>15</sup> Bei Analysen der Fragen zu Begleitwegen von Kindern ist noch ein Indikator zu bilden. Dann sind gegebenenfalls auch noch Entfernungen zu Grundschulen zu ermitteln, die bisher nicht vorliegen.

tanzen in den suburbanen Gebieten vor allem auf die Randwanderer zurückgehen, während die alleingesessenen (autochthonen) Bevölkerungsteile in Suburbia deutlich kürzere Wege zurücklegt, vor allem zur Arbeit, aber auch zum Einkauf oder in der Freizeit (Kasper/Scheiner 2006).

Eine gebildete Variable beschreibt deshalb, ob es sich bei einer Person um einen Randwanderer handelt. Dafür müssen zwei Bedingungen gegeben sein: Die Person muss in einem Umlandgebiet leben und sie muss früher (d.h. hier: seit 1989, da noch frühere Wohnorte nicht erfragt wurden) einmal in Köln gelebt haben.

Dabei handelt es sich um eine Binärvariable, die aufgrund ihrer Verteilung nicht für alle Arten von Analysen geeignet ist. Zudem ist ihre Aussagekraft stark vereinfacht, weil sie keinerlei kleinräumliche Differenzierungen des gegenwärtigen und des früheren Standorts zulässt. Die Bandbreite der raumstrukturellen Indikatoren für die Gebiete zeigt jedoch sehr differenzierte Ausprägungen. So schneiden erwartungsgemäß die Zentren Kerpen und Overath bei der Ausstattung mit Einzelhandel und Dienstleistungen vergleichsweise gut ab. Beim ÖPNV treten in Sindorf und Overath (Schienenanschluss) die besten Werte auf. Die Arbeitsplatzausstattung ist aufgrund des großen Gewerbegebiets in Sindorf am besten, und in der Erreichbarkeit von Bevölkerungspotenzialen schneiden aufgrund der Stadtgröße Kerpens die Gebiete Kerpen und Sindorf am besten ab. Sie übertreffen sogar deutlich das Kölner Stadtrandgebiet Zündorf. Im Folgenden wird deshalb die Methodik der Bildung einer Ordinalvariablen beschrieben, die auf eine genauere Differenzierung des Wanderungstyps abzielt.

Als Ansatzpunkt wird von der Vorstellung eines Stadt-Umland-Kontinuums ausgegangen. Für den gegenwärtigen Wohnort liegen fünf Kategorien vor, die sich vereinfachend auf einer solchen Skala abnehmender Urbanität anordnen lassen: Nippes, Longerich, Zündorf, Kerpen-Stadt/Overath-Stadt, Heiligenhaus/Sindorf. Für den vorherigen Wohnort liegt ebenfalls eine gute Differenzierung vor, sofern dieser in Köln liegt. Im Rahmen des Projekts StadtLeben wurde eine Klassifikation in drei Kategorien vorgenommen: Innenstadt, Innenstadtrand, Stadtrand (beschrieben in Scheiner 2005b, dort Kap. 4.1.2). Dies erlaubt bei Binnenwanderungen innerhalb Kölns eine Aussage über die Richtung der Wanderung.

- So lässt sich beispielsweise eine Wanderung von Köln-Innenstadt nach Longerich als "schwache Randwanderung" charakterisieren, eine Wanderung von Köln-Innenstadt nach Zündorf als "etwas stärkere Randwanderung", eine Wanderung von Köln-Innenstadt nach Sindorf oder Heiligenhaus als "extreme Randwanderung".

Liegt allerdings der vorherige Wohnort außerhalb Kölns, wurde in der Befragung nur die Gemeinde

erhoben. Dies erlaubt keine Differenzierung zwischen Stadt- oder Ortsteilen, wie dies beispielsweise bei der Auswahl der Untersuchungsgebiete vorgenommen wurde, bei denen zwischen den Hauptorten (Kerpen-Stadt, Overath-Stadt) und Ortsteilen (Sindorf, Heiligenhaus) unterschieden wird, um der kleinräumlichen Differenzierung der Siedlungsentwicklung und des Verkehrshandelns Rechnung zu tragen.

- Vereinfachend wird deshalb davon ausgegangen, dass die Angabe "letzter Wohnsitz Kernstadt (in einem Agglomerationsraum oder einem verstärkten Raum)" einem "mittleren" Quartier der Kernstadt entspricht, also etwa einem Quartier am Innenstadtrand. Dies bedeutet, dass der Zuzug aus einer Kernstadt – etwa aus Stuttgart oder Hannover – nach Köln-Nippes als "schwache Stadtwanderung" gewertet wird. Erfolgt der Zuzug nach Köln-Zündorf, dann wird dies als "schwache Randwanderung" gewertet.

Auch bei Wanderungen innerhalb des suburbanen Raums oder aus dem ländlichen Raum in eines der suburbanen Quartiere muss von vereinfachenden Annahmen ausgegangen werden. In der Klassifikation der vorherigen Wohnsitzes wird hier nicht zwischen suburbanen und ländlichen Räumen unterschieden, weil so grobe Unterschiede nicht geeignet sind, durch Wanderungen ausgelöste Veränderungen der Verkehrsnachfrage zu erfassen. Zwar liegt die Gemeindegröße der Herkunftsgemeinde vor, die zur Typologie herangezogen werden könnte. Diese erlaubt jedoch ebenfalls keine adäquate Typisierung der Wanderungsströme, weil es sich bei den betreffenden Herkunftsgemeinden im ländlichen Raum meist um Klein- und Mittelstädte mit 5.000-100.000 Einwohnern handelt, die somit in der gleichen Größenordnung liegen wie die Zuzugsgemeinden Kerpen und Overath. Zudem handelt es sich bei all diesen Wanderungen um Zuzüge aus anderen Regionen, bei denen ohnehin nicht wie bei der intraregionalen Wanderung von Bindungen an den ehemaligen Wohnstandort mit der Folge langer Alltagswege ausgegangen werden kann. Zudem handelt es sich nur um eine geringe Zahl von Fällen (20 Zuzüge aus ländlichen Räumen, 30 aus verstärkten Räumen). Deshalb wird hier auf eine Differenzierung der Herkunftsgemeinde verzichtet. Stattdessen wird eine Differenzierung der Zielorte – also der gegenwärtigen Wohnstandorte – vorgenommen.

- Vereinfachend wird davon ausgegangen, dass ein Zuzug aus einer beliebigen Umlandgemeinde oder ländlichen Gemeinde nach Kerpen-Stadt oder Overath-Stadt weder eine Stadt- noch eine Randwanderung darstellt, sondern der Befragte auf der siedlungsstrukturellen Hierarchie von urban bis ländlich auf dem gleichen Hierarchiestufe bleibt. Erfolgt der Zuzug dagegen nach Sindorf oder Heiligenhaus, wird dies als "schwache Randwanderung" gewertet.

Auf der Grundlage dieser Überlegungen wurde eine ordinalskalierte Variable "Randwanderung" gebildet (Tabelle 14). Diese besitzt acht Ausprägungen zwischen -3 (Wanderung aus suburbanem oder ländlichem Raum nach Nippes) und +4 (Wanderung aus der Kölner Innenstadt in ein Umlandwohnggebiet).

Jetziger Wohnsitz	Letzter Wohnsitz	Köln Innen- stadt	Köln Innen- stadtrand	Köln Stadt- rand	Andere Kern- stadt	Sub- urban / ländlich	keine Wande- rung
Köln Innenstadt (Nippes)		0	-1	-2	-1	-3	0
Köln Innenstadtrand (Longerich)		1	0	-1	0	-2	0
Köln Stadtrand (Zündorf)		2	1	0	1	-1	0
Umland Zentrum (Kerpen, Overath)		3	2	1	2	0	0
Umland Wohngebiet (Sindorf, Heiligenhaus)		4	3	2	3	1	0

**Tabelle 14: Variable "Klassifikation von Wanderungen"**

Die Tabelle gibt die Ausprägungen einer Variablen an, in der Wanderungen nach dem Grad der Veränderung der "Urbanität des Wohnstandortes" durch die Wanderung klassifiziert werden. Positive Ausprägungen sind Randwanderungen, negative Ausprägungen sind Stadtwanderungen (Zunahme von Urbanität).

Bei der Verwendung dieser Variablen für Verkehrsmodelle ist allerdings dem Problem der Asymmetrie von Effekten Rechnung zu tragen. Für die Erklärung der Verkehrsnachfrage steht hinter dieser Variablen die implizite These, dass Stadtwanderungen gegenüber Randwanderungen spiegelbildliche Effekte besitzen. Wenn also beispielsweise Randwanderer höhere Distanzen zurücklegen und mehr Auto fahren als sesshafte Suburbaniten, dann müssten nach diesem Modell Stadtwanderer geringere Distanzen zurücklegen und weniger Auto fahren als sesshafte Städter. Diese Annahme ist aber nicht sinnvoll. Deskriptive Analysen deuten darauf hin, dass Zuwanderer in die Stadt eher *längere* Wege zurücklegen und *mehr* Auto fahren als sesshafte Städter, auch wenn demografische Unterschiede kontrolliert werden. Dies ist auch plausibel, weil Stadtwanderungen zwar mit einer Verringerung der Motorisierung einhergehen (Scheiner 2005a). Die Abschaffung einiger Pkw durch Stadtwanderer nach der Wanderung ist aber vor dem Hintergrund eines relativ hohen Motorisierungsniveaus vor der Wanderung zu sehen und führt keineswegs zu einer geringeren Motorisierung von Stadtwanderern gegenüber sesshaften Städtern. Zudem ist aufgrund der Bindungen an den vorherigen Wohnstandort auch bei Stadtwanderern eher von längeren Wegen auszugehen als bei sesshaften Städtern. Diese Problematik muss bei der Anwendung der Variablen gut reflektiert werden.

#### 4.9 Verkehrsnachfrage

Für die Erhebung von Verkehrshandeln gibt es im Wesentlichen zwei Methoden. In Wegetagebüchern (travel diaries) werden für einen oder mehrere vorgegebene Stichtage alle Wege mit ihren Merkmalen (Start- und Zielort, Start- und Zielzeit, Verkehrsmittel, Entfernung, Zweck, Begleitung) aufgezeichnet. Dies hat den Vorteil relativ großer Genauigkeit, aber den

Nachteil, dass sich die Daten nur auf die vorgegebenen Tage (in der Regel nur ein Tag) beziehen. Aus solchen Daten können keine individuellen Handlungsmuster abgeleitet werden, weil über den Stichtag hinaus das Handeln unbekannt ist.

Bei der frequent activities-Methode dagegen, die in StadtLeben angewandt wurde, werden für ausgewählte Aktivitäten retrospektiv 'übliche' Handlungsweisen erfasst: geschätzte Aktivitätshäufigkeit, üblicher Aktivitätssort mit Entfernung, übliches Verkehrsmittel, evtl. Begleitung, Wegedauer). Dies hat den Nachteil geringerer Genauigkeit, aber den Vorteil, dass sich daraus ein, wenn auch generalisiertes, individuelles Bild des Handelns auf der Individual-ebene zeichnen lässt. Erhoben wurden in StadtLeben die Aktivitäten Arbeiten, Ausbildung, Einkauf für den täglichen Bedarf, Großeinkauf, Shopping, Behörden- und Verwaltungsgänge, private Erledigungen, private Besuche, aktiver Sport, Gastronomiebesuche, Kulturveranstaltungen, Sportveranstaltungen, Disco/Konzert, Spaziergänge, Ausflüge.

Die Angaben der Befragten wurden zu aktivitäts- und verkehrsmittelspezifischen Häufigkeiten und Entfernungen zusammengefasst. Im Ergebnis liegen für jede erfragte Aktivität sowie summarisch für Freizeit-, Versorgungs- und Pflichtaktivitäten (Arbeit, Ausbildung) Aktivitätshäufigkeiten je Woche, zurückgelegte Distanzen je Woche und zurückgelegte Distanzen je Weg vor. Diese sind getrennt für einzelne Verkehrsmittel sowie über alle Verkehrsmittel hinweg ermittelt worden. Die Aktivitätshäufigkeit entspricht grob der in den Verkehrswissenschaften üblicheren Wegehäufigkeit, wenn Verzerrungen durch Wegekopplungen sowie Rückwege nach Hause unberücksichtigt bleiben.

Untersucht werden hier schwerpunktmäßig Arbeits-, Versorgungs- und Freizeitwege. In die Versorgungs-

wege gehen die Angaben zu folgenden Aktivitäten ein: Einkauf für den täglichen Bedarf, Großeinkauf, Shopping, Behörden- und Verwaltungsgänge, private Erledigungen. Summarisch sind diese im Wesentlichen durch Einkäufe geprägt. Allein 68% entfallen auf den kleinen Einkauf. Nimmt man Großeinkäufe und Shopping hinzu, umfassen Einkäufe zusammen 88% der genannten Versorgungsaktivitäten. Freizeitwege umfassen private Besuche, aktiver Sport, Gastronomiebesuche, Kulturveranstaltungen, Sportveranstaltungen, Disco/Konzert, Spaziergänge und Ausflüge.

Im Rahmen des Projekts wird das Verkehrshandeln sowohl unter dem Aspekt Aktivitätshäufigkeit als auch anhand von Wegelängen (Distanzen) und Verkehrsmittelnutzung untersucht. Diese Indikatoren sind auch in Kombination miteinander analysierbar. Beispielsweise hängen Wegelänge und Verkehrsmittelnutzung eng miteinander zusammen. So kommen nicht-motorisierte Verkehrsmittel nur für relativ kurze Wege in Frage. Die Verkehrsmittelnutzung kann also in Abhängigkeit von der Wegelänge modelliert werden (oder umgekehrt!).

Die zurückgelegten Distanzen besitzen sehr schiefe Verteilungen, die sich durch natürliche Logarithmierung stark mindern lässt. Die Distanzangabe 'null' wird dabei durch 0.001 km ersetzt. Dabei handelt es sich nicht um fehlende Werte, sondern um Aktivitäten am Wohnstandort (z.B. Arbeit im eigenen Haus).

In ergänzenden Analysen werden auch Indikatoren räumlicher Orientierungen verwendet. So hat sich im Projekt StadtLeben gezeigt, dass die langen alltäglichen Distanzen von Randwanderern teilweise auf deren Rückorientierungen in die Kernstadt zurückzuführen sind (Kasper/Scheiner 2006). Deshalb wird in weitergehenden Analysen auch berücksichtigt, ob eine Person einen Arbeits- oder Ausbildungsplatz in Köln besitzt.

Die Pkw-Verfügbarkeit schließlich wird als ordinalskalierte Variable mit vier Ausprägungen in den Analysen verwendet:

- kein Pkw im Haushalt;
- Pkw im Haushalt der dem/der Befragten nicht zur Verfügung steht;
- Führerschein und Pkw im Haushalt der dem/der Befragten zeitweise zur Verfügung steht;
- Führerschein und Pkw im Haushalt der dem/der Befragten ständig zur Verfügung steht.

Diese Variable dürfte dann als metrisch skaliert interpretiert werden, wenn man gleiche Abstände zwischen den vier Ausprägungen annehmen dürfte. Die tatsächliche Pkw-Nutzung in den vier Gruppen legt nahe, dass dies zumindest approximativ gilt (Tabelle 15). Keine zwei Gruppen sind extrem nahe zusammen oder extrem weit auseinander.

Pkw-Verfügbarkeit	Pkw-Nutzung je Woche		n
	Distanz	Häufigkeit	
kein Pkw im Haushalt	12,2	0,8	305
Pkw im Haushalt...			
...nicht verfügbar	34,9	2,2	88
...zeitweise verfügbar	67,0	4,8	273
...ständig verfügbar	104,3	6,7	1454
Alle	83,2	5,4	2120

**Tabelle 15: Pkw-Nutzung nach Pkw-Verfügbarkeit (Mittelwerte)**

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

Alternativ wird dennoch auch mit der Anzahl der Pkw je erwachsener Person im Haushalt gearbeitet, da es sich dabei um einen metrisch skalierten Indikator handelt. Dieser spiegelt allerdings nicht notwendigerweise die Pkw-Verfügbarkeit eines Individuums im Haushalt wider, weil hohe Werte auch bei Personen auftreten können, die selten über einen Pkw verfügen und umgekehrt.

## 5 Methodische Analysen: Evaluation der raumstrukturellen Indikatoren

### 5.1 Lohnt die personenbezogene Differenzierung der Wohnstandorte innerhalb der Gebiete?

Aus den Luftliniendistanzen zwischen dem Wohnstandort der Befragten und den Standorten der kartierten Einzelhandels-, Dienstleistungs-, Ausbildungs- und Freizeiteinrichtungen wurden personenbezogene Indikatoren der Erreichbarkeit gebildet. Danach stellt sich die Frage, ob sich dies insoweit lohnt, als innerhalb eines Untersuchungsgebietes nennenswerte Differenzierungen entstehen. Die Korrelationen zwischen den personenbezogenen Werten der Indikatoren und den jeweiligen Mittelwerten für die Untersuchungsgebiete liegen zwischen  $r=0.84$  (Freizeiteinrichtungen ohne Gastronomie) und  $r=0.95$  (Lebensmittelgeschäfte ohne Super-/Verbrauchermärkte). Auf den ersten Blick scheint also die Differenzierung nur mäßig zu sein.

Die Spannweite der Indikatoren innerhalb der Gebiete ist allerdings teilweise erheblich (Tabelle 16). Beispielsweise variiert die Distanz zum nächsten Supermarkt in Longerich zwischen 50 m und 1,11 km – eine Distanz, die von vielen Nachfragern bereits mit dem Pkw zurückgelegt wird. Die Anzahl der Lebensmittelgeschäfte (ohne Super- und Verbrauchermärkte) im Radius von 650 m Luftlinie variiert für Befragte in Kerpen Stadt zwischen 0 und 32; sicher kein unerheblicher Unterschied.

Werden die Werte für jedes Gebiet danach differenziert, ob die befragte Person eher zentral (max. 250 m vom Mittelpunkt des Gebiets entfernt) oder eher peripher im Untersuchungsgebiet lebt (Tabelle 17), zeigt sich erwartungsgemäß, dass meist die Befragten mit zentralem Wohnstandort in einem Gebiet ein breiteres Spektrum von Einrichtungen kleinräumlich erreichen können.

Diese Differenzierungen sagen noch nichts darüber aus, ob von ihnen Effekte auf die Verkehrsnachfrage

ausgehen, und wenn ja, welche. Dies erscheint angesichts der kleinräumlichen Unterschiede im Verkehrshandeln (Holz-Rau 1991) allerdings wahrscheinlich. Zumindest sind sie eine notwendige Voraussetzung für die Differenzierung des Verkehrshandelns innerhalb der Gebiete. Inwieweit dies zu relevanten substantziellen Ergebnissen führt, ist im weiteren Projektverlauf zu prüfen.

		Distanz		Anzahl... in 650m Radius Luftlinie					
		nächster Super-/Verbrauchermarkt (Mittelwert, km)	Super-/Verbrauchermärkte	Andere Lebensmittelgeschäfte	Andere Geschäfte	Dienstleistungen (ohne Medizin)	Dienstleistung medizin.	Freizeiteinrichtungen (ohne Gastronomie)	Gastro- nomie
Ehrenfeld	Mittelwert	0,18	6	67	148	82	67	76	103
	Minimum	0,01	3	38	85	45	26	40	60
	Maximum	0,41	8	83	176	104	89	90	120
Nippes	Mittelwert	0,18	7	67	91	64	38	92	129
	Minimum	0,00	2	22	19	16	15	45	34
	Maximum	0,46	10	87	114	81	50	121	166
Stammheim	Mittelwert	0,25	2	14	11	9	8	66	7
	Minimum	0,03	1	10	5	4	0	56	3
	Maximum	0,53	2	16	14	12	11	70	10
Longerich	Mittelwert	0,42	2	9	17	13	13	8	17
	Minimum	0,05	0	5	7	7	7	3	8
	Maximum	1,11	2	12	27	19	18	10	23
Esch	Mittelwert	1,66	0	0	0	0	0	37	2
	Minimum	1,23	0	0	0	0	0	7	1
	Maximum	2,11	0	0	0	0	0	49	3
Zündorf	Mittelwert	0,39	1	8	13	16	11	44	28
	Minimum	0,04	0	0	1	1	2	5	2
	Maximum	1,05	2	13	19	24	16	69	47
Kerpen Stadt	Mittelwert	0,42	2	19	51	27	15	53	39
	Minimum	0,02	0	0	1	1	1	3	1
	Maximum	0,97	5	32	81	45	26	85	63
Sindorf	Mittelwert	0,43	2	15	20	8	10	40	16
	Minimum	0,01	0	4	2	0	0	6	0
	Maximum	0,92	3	28	33	17	17	74	34
Overath Stadt	Mittelwert	0,44	4	11	29	19	20	27	14
	Minimum	0,00	0	0	0	0	0	0	0
	Maximum	1,05	7	17	47	30	30	52	22
Heiligenhaus	Mittelwert	2,29	0	5	3	2	5	18	3
	Minimum	1,27	0	0	0	0	0	0	0
	Maximum	3,50	0	9	5	5	9	39	6
Insgesamt	Mittelwert	0,66	3	21	36	23	18	44	36
	Minimum	0,00	0	0	0	0	0	0	0
	Maximum	3,50	10	87	176	104	89	121	166

**Tabelle 16: Vergleich zwischen Gebieten: Mittelwerte und Spannweite der Indikatoren für Freizeit- und Versorgungsangebote innerhalb der Gebiete**

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

	Lage Wohnstandort innerhalb des UG	Distanz nächster Super/Verbrauchermarkt (Mittelwert, km)	Anzahl... in 650m Radius Luftlinie						
			Super-/Verbrauchermärkte	Andere Lebensmittelgeschäfte	Andere Geschäfte	Dienstleistungen (ohne Medizin)	medizin. Dienstleistung	Freizeiteinrichtungen (ohne Gastronomie)	Gastro- nomie
Ehrenfeld	peripher*	0,20	5,3	61,0	135,9	75,4	58,4	69,0	92,5
	zentral*	0,16	7,0	72,5	158,7	87,8	73,9	81,2	111,0
Nippes	peripher*	0,21	7,4	62,3	84,5	58,6	34,8	86,0	118,5
	zentral*	0,11	7,5	77,4	106,1	76,1	44,6	106,8	156,0
Stammheim	peripher*	0,37	1,5	12,9	8,9	7,7	5,1	65,5	6,1
	zentral*	0,21	1,9	14,2	11,1	9,2	8,6	66,7	7,5
Longerich	peripher*	0,49	1,6	9,2	17,2	13,7	12,4	7,9	16,2
	zentral*	0,17	2,0	9,3	15,6	11,4	13,4	7,5	18,1
Esch	peripher*	1,70	0,0	0,0	0,0	0,0	0,0	32,0	2,1
	zentral*	1,59	0,0	0,0	0,0	0,0	0,0	46,8	3,0
Zündorf	peripher*	0,44	1,2	7,5	12,0	14,6	10,0	40,1	25,4
	zentral*	0,22	1,7	11,9	17,0	22,4	15,0	59,0	39,1
Kerpen Stadt	peripher*	0,45	2,0	17,8	48,2	25,7	14,0	50,3	36,6
	zentral*	0,17	3,0	29,3	76,6	42,3	24,7	79,5	60,2
Sindorf	peripher*	0,44	1,4	14,5	18,7	8,0	9,2	38,4	16,0
	zentral*	0,41	2,1	19,7	23,6	9,0	13,2	44,4	16,7
Overath Stadt	peripher*	0,48	3,7	10,0	26,2	17,0	18,0	25,2	12,8
	zentral*	0,28	6,2	17,0	42,8	28,5	29,4	38,4	22,0
Heiligenhaus	peripher*	2,29	0,0	4,3	2,2	2,0	4,1	15,8	2,8
	zentral*	2,26	0,0	8,2	4,8	3,1	7,8	37,4	4,5
Insgesamt	peripher*	0,74	2,3	17,7	31,3	20,1	15,0	38,5	31,0
	zentral*	0,44	3,4	29,8	50,6	32,3	25,6	61,2	48,8

**Tabelle 17: Vergleich der Erreichbarkeit von Freizeit- und Versorgungsangeboten zwischen zentraler und peripherer Lage innerhalb der Gebiete**

\*peripher:  $\geq 250$ m von Mitte des UG

\*zentral:  $< 250$ m von Mitte des UG

Quelle: eigene Analysen. Daten: Projekt StadtLeben.

## 5.2 Lohnt der Aufwand der Bildung eines Indikators 'Arbeitsplatzerreichbarkeit'?

Der hohe Aufwand einer kleinräumlichen Ermittlung der Erreichbarkeit von Arbeitsplätzen legt für zukünftige Forschungen die Frage nahe, ob es möglich gewesen wäre, mit deutlich geringerem Aufwand einen ähnlich guten Indikator der Arbeitsplatzerreichbarkeit zu konstruieren. Dies ist umso mehr relevant, als die kleinräumliche Arbeitsplatzverteilung nur durch die Initiative eines Mitarbeiters der Firma PTV überhaupt verfügbar war und für zukünftige Forschungen damit nicht gerechnet werden kann. Die Arbeitsplatzzentralität ist eng mit anderen Zentralitätsindikatoren verbunden, so dass möglicherweise ein allgemeiner Indikator "Urbanität" oder "Zentralität" die Arbeitsplatzerreichbarkeit gut abbilden könnte, ohne dass Informationen über die räumliche Verteilung der Arbeitsplätze vorliegen. Der entsprechende Indikator könnte beispielsweise In-

formationen über die Einzelhandelsausstattung oder die Bevölkerungsdichte enthalten.

In der Tat zeigen sich zwischen der Arbeitsplatzerreichbarkeit und anderen Indikatoren zentralörtlicher Ausstattung starke korrelative Zusammenhänge.

Die höchste Korrelation zur Arbeitsplatzerreichbarkeit erzielen die beiden Indikatoren für die Bevölkerung, nämlich die Bevölkerungsdichte im Gebiet sowie die mit einem Gravitationsansatz ermittelte Bevölkerungserreichbarkeit. Aber auch die Ausstattung mit Freizeiteinrichtungen, Einzelhandel und Dienstleistungen, die ÖPNV-Qualität sowie – besonders einfach zu ermitteln – die Entfernung zu Köln sind stark mit der Arbeitsplatzerreichbarkeit korreliert. Die hohen Korrelationen sollten allerdings nicht darüber hinwegtäuschen, dass im Einzelnen deutliche Unterschiede zwischen den Indikatoren auftreten. So erreicht Sindorf beim Versorgungsangebot einen Wert von  $-0.40$ , bei der Arbeitsplatzerreichbarkeit einen Wert von  $+0.61$  (standardisierte Werte).

	Arbeitsplatzerreichbarkeit ( $\alpha=1$ )	Arbeitsplatzdichte (je qkm) im Gebiet	Freizeitangebot*	Einzelhandel + Dienstleistungen*	Erreichbarkeit Bevölkerung	Bevölkerungsdichte	ÖPNV-Systemqualität	Entfernung nach Köln- Zentrum
Arbeitsplatzerreichbarkeit ( $\alpha=2$ )	0,20	-0,10	-0,02	-0,02	0,08	0,09	0,09	0,48
Arbeitsplatzerreichbarkeit ( $\alpha=1$ )	1,00	0,78	0,76	0,75	0,95	0,90	0,76	-0,73
Arbeitsplatzdichte (je qkm) im Gebiet	0,78	1,00	0,85	0,92	0,80	0,93	0,53	-0,67
Freizeitangebot*	0,76	0,85	1,00	0,91	0,77	0,80	0,49	-0,60
Einzelhandel + Dienstleistungen*	0,75	0,92	0,91	1,00	0,74	0,84	0,63	-0,53
Erreichbarkeit Bevölkerung	0,95	0,80	0,77	0,74	1,00	0,88	0,69	-0,77
Bevölkerungsdichte	0,90	0,93	0,80	0,84	0,88	1,00	0,60	-0,68
ÖPNV-Systemqualität	0,76	0,53	0,49	0,63	0,69	0,60	1,00	-0,53
Entfernung nach Köln-Zentrum	-0,73	-0,67	-0,60	-0,53	-0,77	-0,68	-0,53	1,00

**Tabelle 18: Korrelationen zwischen Arbeitsplatzerreichbarkeit und anderen raumstrukturellen Merkmalen der Untersuchungsgebiete**

\* Anzahl der Angebote im 650m-Radius, Mittelwert pro Gebiet aus personenbezogenen Werten

Korrelationen für n=10 Gebiete

Quelle: eigene Analysen.

Die Zusammenhänge bestehen nur dann, wenn für die Arbeitsplatzerreichbarkeit ein Raumwiderstandskoeffizient von  $\alpha=1$  verwendet wird (vgl. Kapitel 4.7.1). Bei höheren Koeffizienten sind die Korrelationen deutlich schwächer. Dies ist auf den ersten Blick überraschend, denn alle verwendeten Proxy-Indikatoren beziehen sich auf Merkmale der Untersuchungsgebiete selbst, und die Arbeitsplatzerreichbarkeit müsste diesen Merkmalen umso mehr ähneln, je höher der Raumwiderstandskoeffizient ist, weil bei hohem Raumwiderstand die Werte stärker von der Arbeitsplatzausstattung des Gebiets selbst und seiner nächsten Umgebung abhängen. Die höchsten Korrelationen sollten also zwischen Arbeitsplatzerreichbarkeit mit Raumwiderstand  $\alpha=2$  und Arbeitsplatzdichte im Untersuchungsgebiet auftreten. Tatsächlich ist diese Korrelation aber mit  $r=-0,10$  schwach negativ. Die Korrelationen der Arbeitsplatzerreichbarkeit mit Raumwiderstand  $\alpha=2$  mit anderen Proxy-Indikatoren sind ebenfalls schwach; die Korrelation mit der Entfernung von Köln ist nun entgegen der Erwartung deutlich positiv, d.h. je weiter von Köln entfernt ein Gebiet ist, desto besser erscheint dort die Arbeitsplatzerreichbarkeit. Diese widersinnigen Ergebnisse sind durch den Ausreißer Sindorf zu erklären, der bei hohen Raumwiderstandskoeffizienten das Verhältnis zwischen den Gebieten stark verzerrt. Dies bedeutet aber auch, dass die Arbeitsplatzausstattung in den Untersuchungsgebieten nicht sehr stark mit den verwendeten Proxy-Indikatoren korreliert. Hohe Korrelationen entstehen erst, wenn die Arbeitsplatzausstattung des näheren Umfeldes, etwa der angrenzenden Ortsteile, berücksichtigt wird. Der Schluss von der Ausstattung eines Gebiets mit Ver-

sorgungseinrichtungen auf die Ausstattung des Gebiets mit Arbeitsplätzen ist also nicht möglich, wohl aber (in den hier vorliegenden Daten) der Schluss von der Ausstattung eines Gebiets mit Versorgungseinrichtungen auf die Ausstattung der weiteren Umgebung mit Arbeitsplätzen. Der Zusammenhang zwischen Arbeitsplatzerreichbarkeit und Gebietsausstattung mit weiteren Wohnfolgeeinrichtungen besteht also eher auf einer generelleren Klassifikation der Lage des Gebiets in der Region.

Die Korrelationen lassen den Schluss zu, dass die aufwändige Ermittlung der Arbeitsplatzerreichbarkeit tatsächlich verzichtbar gewesen wäre – allerdings unter der unrealistischen Prämisse, dass die Ergebnisse der hier vorgestellten Korrelationsanalyse bereits bekannt gewesen wären.

Für zukünftige Forschungen sind dabei mindestens drei Überlegungen zu berücksichtigen:

1. Die hier ermittelten Korrelationen gelten für eine Auswahl von zehn diskreten, sich stark unterscheidenden Gebieten. Würde man eine ähnliche Analyse für eine größere Zahl von Gebieten mit geringerer Differenzierung durchführen – im Extremfall für eine in statistische Einheiten unterteilte Gesamtregion – dann würden sich vermutlich andere, komplexere Überlagerungen verschiedener Gebietsmerkmale ergeben (z.B. gute Einzelhandelsausstattung bei schlechter Arbeitsplatzerreichbarkeit). Dann dürften die Korrelationen zwischen den verschiedenen Raumstrukturmerkmalen geringer ausfallen.

2. Die Ausstattung eines kleineren Untersuchungsgebietes mit Einzelhandel, Dienstleistungen oder Freizeiteinrichtungen sowie die Bevölkerungsdichte können als Proxy-Indikatoren für „Urbanität“ dienen, die wiederum die Arbeitsplatzausstattung im näheren Umfeld, aber nicht notwendigerweise in dem Gebiet selbst gut beschreibt.
3. Eine adäquate Abbildung des Arbeitsplatzangebots hängt stark von den Details der Indikatorbildung ab. Beispielsweise kann die Höhe des Raumwiderstandskoeffizienten  $\alpha$  nicht a priori festgelegt werden, sondern die mit einem bestimmten Raumwiderstandskoeffizienten erzielten Ergebnisse sind auf Plausibilität zu prüfen (etwa mit Hilfe anderer Urbanitätsindikatoren) und durch Sensitivitätsanalysen zu stützen. Dies bedeutet aber auch, dass die Substitution eines Arbeitsplatzindikators durch Proxy-Variablen nur möglich ist, wenn geeignete Arbeitsplatzindikatoren bereits vorliegen und die Zusammenhänge mit den Proxy-Variablen geprüft sind. Dann besteht aber für diese Substitution kein Anlass.

Ein Rückgriff auf Proxy-Indikatoren zur Abbildung der Arbeitsplatzerreichbarkeit dürfte also nur dann adäquat sein, wenn drei Bedingungen gleichzeitig gegeben sind:

- Die Untersuchungsgebiete sind so typisierend ausgewählt, dass sie sich in verschiedenen Zentralitätsindikatoren gleichermaßen deutlich unterscheiden. Eine solche Annahme ist sorgfältig zu begründen und beispielsweise durch qualitative Einschätzungen lokaler Experten zu stützen.
- Der Indikator der Arbeitsplatzerreichbarkeit hat nicht zum Ziel, die Ausstattung mit Arbeitsplätzen im mikroräumlichen Kontext abzubilden. Auf mikroräumlicher Ebene (z.B. innerhalb eines Ortsteils) kann die Abbildung durch Proxy-Variablen sehr schlecht ausfallen.
- Ziel ist nicht die Abbildung des absoluten Niveaus der Arbeitsplatzausstattung, sondern lediglich die Abbildung relativer Niveauunterschiede zwischen Gebieten.

## 6 Analyseverfahren

Bisher hat sich dieses Papier auf die Aspekte Datenerhebung und Datenaufbereitung konzentriert. Abschließend wird ein Ausblick auf die Datenanalyse gegeben.

Die in Kapitel 2 diskutierten Zusammenhänge können mit Hilfe linearer Strukturgleichungsmodelle untersucht werden. Diese werden in der Verkehrsfor- schung in jüngerer Zeit zunehmend angewandt (Golob 2001). Verbreitet sind sie insbesondere in der Mobilitätspsychologie, vor allem in Anwendungen der Theorie des geplanten Verhaltens nach Ajzen (1991). Dabei wird speziell die Verkehrsmittel- nutzung auf subjektive Normen, Einstellungen und

die wahrgenommene Verhaltenskontrolle zurückge- führt (Bamberg 2004). In mehr verkehrswissenschaft- lich ausgerichteten Modellen wird die Verkehrsmittel- verfügbarkeit als intervenierende Größe betrachtet, indem sie einerseits die Verkehrsmittelnutzung und andere Aspekte des Verkehrshandelns (z.B. die zu- rückgelegten Distanzen) "erklärt", andererseits aber ihrerseits bereits Teil der Entscheidung für ein be- stimmtes Verkehrshandeln ist. Diese Entscheidung ist abhängig von soziodemografischen Faktoren der Lebenslage (Simm/ Axhausen 2001).

### 6.1 Methodik linearer Strukturgleichungsmodelle

Lineare Strukturgleichungsmodelle sind eine jüngere Weiterentwicklung der Pfadanalyse, die auch als Kausalanalyse bekannt ist, in Kombination mit der Faktorenanalyse.

Die Pfadanalyse stellt eine verallgemeinerte Form der Regressionsanalyse dar, die sich nicht auf eine Gleichung beschränkt, sondern mehrere Gleichungen simultan zu lösen versucht. Gegenüber der klassi- schen Regressionsanalyse mit einer Gleichung hat dies den Vorteil, dass komplexe Beziehungen zwi- schen Variablen sehr flexibel modelliert werden kön- nen. Insbesondere können intervenierende Variablen modelliert werden, die einerseits eine abhängige Funktion einnehmen und andererseits für andere Variablen unabhängig (erklärend) sind. Damit lassen sich direkte und indirekte Effekte einer Variablen auf eine andere Variable unterscheiden. Indirekte Effekte sind moderiert über intervenierende Variablen. Dies erlaubt es je nach Substanz und Zielsetzung des Modells auch, politische Konzepte zu identifizieren, die auf eine erwünschte Wirkung (abhängige Vari- able) einen direkten oder einen indirekten Effekt besitzen. Der pfadanalytische Teil eines Strukturglei- chungsmodells stellt das sogenannte Strukturmodell dar, das den inhaltlichen Kern bildet.

Über die Strukturmodelle hinaus kann mit Messmo- dellien gearbeitet werden. Diese stellen die faktoren- analytischen Teile des Strukturgleichungsmodells dar. Die Faktoren werden hier als latente (nicht beobachtete) Variablen bezeichnet, deren Ausprägungen im Rahmen eines Messmodells aus anderen (beobachteten) Variablen geschätzt werden (Bollen 1989). In der Arbeit mit latenten Variablen wird Abstand genommen von der Annahme fehlerfreier Messungen. Damit wird neben dem Modellieren intervenierender Variablen eine weitere Qualität linearer Strukturgleichungsmodelle gegenüber der 'klassischen' Regressionsanalyse ausgeschöpft. Die Annahme fehlerfreier Messungen ist insbesondere bei komplexen, abstrakten Begriffen wie Lebenslage oder Lebensstil nicht realistisch, weil nicht davon ausgegangen werden kann, dass solche Begriffe direkt messbar sind. Die Verwendung latenter Vari- ablen geht davon aus, dass die gemessenen Vari-



ablen – bei der Lebenslage etwa Bildungsniveau, Einkommen oder Haushaltstyp – lediglich mehr oder weniger gute Repräsentationen des eigentlich interessierenden latenten Konstrukts Lebenslage sind. Die latente Variable wird also als Realität betrachtet, die gemessenen Variablen dagegen nur als 'unscharfe' und damit fehlerhafte Indikatoren derselben. Im Modell sind die gemessenen Variablen konsequenterweise als abhängig von der latenten Variable gesetzt.

Auch wenn man Strukturgleichungsmodelle auf der Grundlage von Individualdaten anwendet, handelt es sich im Grunde um aggregierte Modelle, die auf Stichprobenmomenten beruhen, d.h. auf Mittelwerten, Varianzen und Kovarianzmatrizen. Die Rohdaten – bei Befragungen also die Angaben einzelner Befragter – sind als Input für die Modelle nicht zwingend notwendig.

Die rechnerische Vorgehensweise besteht nun darin, dass zunächst die Kovarianzmatrix zwischen den Variablen berechnet wird. Die Parameter werden dann so berechnet, dass die empirische Kovarianzmatrix möglichst gut reproduziert wird. Die Abweichung zwischen empirischer und modellierter Kovarianzmatrix wird auf Signifikanz geprüft. Der Signifikanztest setzt die multivariate Normalverteilung der beobachteten Variablen voraus. Die Modellanpassung an die Daten kann dann als gut angesehen werden, wenn der Test *nicht* signifikant ausfällt. Aus großen Stichproben ergibt sich das Problem, dass die Modelltests "zu gut" sind und auch geringe Abweichungen zwischen Modell und Datenstruktur signifikant ausfallen. Zur Bewertung der Modellgüte gibt es aber verschiedene heuristische Indizes, für die belegte Entscheidungsregeln existieren (Golob 2001, S. 10ff, Wheaton 1988).

Anschließend wird die Signifikanz jedes einzelnen Parameters getestet. Dafür wird der entsprechende Parameter gleich Null gesetzt und das zugehörige Strukturgleichungsmodell neu geschätzt. Auf Signifikanz geprüft wird dann der Unterschied zwischen den modellierten Kovarianzmatrizen mit dem frei geschätzten und dem auf Null gesetzten Parameter. Die Parameter werden meist mit der Maximum-Likelihood-Methode (ML) ermittelt.

Für nicht normalverteilte Daten hat Browne (1984) ein asymptotisch verteilungsfreies Schätzverfahren (ADF; asymptotically distribution free) entwickelt. Dies kann auch bei ordinal skalierten Variablen verwendet werden. Dabei werden Matrizen polychorischer Korrelationen an Stelle von Kovarianzmatrizen verwendet.

Nach Simulationsstudien gilt allerdings das ML-Schätzverfahren bei größeren Stichproben als robust gegen Verletzungen der Verteilungsvoraussetzung (Golob 2001, S. 9) und zumindest bei nicht normalverteilten *stetigen* Variablen als besser als das ADF-Verfahren (Schermele-Engel/Moosbrugger/Müller

2003, S. 27). Bentler und Chou (1987) finden keine Hinweise, dass asymptotisch verteilungsfreie Verfahren, selbst wenn sie aus theoretischer Perspektive für kategoriale Variablen angemessener sein sollten als die ML-Schätzung, in der praktischen Anwendung besser abschneiden. Sie schlagen vor, Ordinalvariablen mit vier oder mehr Ausprägungen als metrisch zu behandeln. Nach Hoogland und Boomsma (1998) schätzt das ADF-Verfahren Standardfehler besser bei einer durchschnittlichen (!) Kurtosis der beobachteten Variablen  $>3.0$  und  $n>400$ . Kühnel (2006, S. 131) fordert für das ADF-Verfahren einen Stichprobenumfang von  $n=500$ , besser  $n=1.000$ . Hoogland und Boomsma (1998) nennen als notwendigen Stichprobenumfang  $n>5*df$ , weil ansonsten die ML Chi-Quadrat Statistik das Modell zu oft zurückweist. Bei einer durchschnittlichen Kurtosis  $>5.0$  und ML-Schätzung sollte  $n>10*df$  sein, beim ADF-Verfahren gilt sogar  $n>20*df$ .

Als Anhaltswert für lineare Strukturgleichungsmodelle im Allgemeinen werden etwa 15 Fälle pro gemessener Variable oder fünf (bei nicht-normalen Daten zehn) Fälle je zu schätzendem Parameter (Bentler/Chou 1987) genannt. Nachdem normalerweise mit jeder gemessenen Variable wenigstens ein Pfadkoeffizient assoziiert ist sowie eine Varianz und ein Fehlerterm zu schätzen ist, kann der Wert von 15 Fällen je gemessener Variable als grober Anhaltspunkt dienen (University of Texas w/o year, Stevens 1996). Darüber hinaus kann mit einer Stichprobengröße von  $n>500$  und mit einer Anzahl an Freiheitsgraden von  $df>30$  eine hohe Power ( $>0.95$ ) beim Hypothesentest erzielt werden, auch wenn Normalverteilung nicht gegeben ist (MacCallum et al. 1996). Die statistische Power entspricht der Wahrscheinlichkeit, dass ein Test eine falsche Nullhypothese korrekterweise ablehnt ( $1-\beta$ ).

Frühere Erfahrungen mit Strukturgleichungsmodellen des Verkehrshandelns (Simma 2000) zeigen, dass Modelle mit latenten Variablen häufig zu äußerst geringer Modellgüte führen. Erste Erfahrungen mit den hier untersuchten Daten gehen in eine ähnliche Richtung. In verkehrswissenschaftlichen Anwendungen werden deshalb manchmal pauschal alle Fehlerkovarianzen zugelassen (z.B. Cao/Mokhtarian/Handy 2007), um die Modellgüte zu erhöhen. Dies geht allerdings auf Kosten eines theoriegeleiteten Vorgehens.

Eine Alternative ist die Arbeit mit reinen Pfadmodellen ohne Einschluss latenter Variablen. Die Arbeit mit Faktorwerten als 'gemessenen Variablen' unter Ausklammerung von Messmodellen ist deshalb in der Verkehrsforschung nicht unüblich (z.B. Cao/Mokhtarian/Handy 2007, Scheiner/Holz-Rau 2007). Der Ausschluss von Messmodellen entspricht der Annahme, dass die verwendeten Variablen fehlerfreie Messungen der interessierenden latenten Konstrukte sind. Tatsächliche Messfehler werden (fälschlicherweise) aus den Modellen heraus gezwungen. Dies ist theoretisch unbefriedigend. Darüber hinaus führt es

zu geringeren Varianzaufklärungen, weil diese von der Validität der Indikatoren abhängen.

Hier wird deshalb mit latenten Variablen gearbeitet. Zur Modellanpassung und zur Validierung der Ergebnisse wird dabei ein recht rigores Verfahren angewandt, das weiter unten beschrieben wird.

Die Analysen werden mit dem Programm AMOS (Analysis of Moment Structures) in den Versionen 5.0 und 6.0 durchgeführt. Dieses ist mit einer benutzerfreundlichen Grafikoberfläche im Programmpaket SPSS integriert verfügbar. Für einige Schritte wird auch das Programm LISREL 8.0 verwendet.

## 6.2 Verteilung der Variablen und Skalierung

Die zu schätzenden Modellen basieren, wie dies in sozialwissenschaftlichen Daten in der Regel der Fall ist, sowohl auf schief verteilten stetigen als auch auf ordinal skalierten und binären Variablen. Die Verwendung binärer Variablen als exogene Variablen ist unproblematisch, jedoch nicht als endogene Variablen.

Das Arbeiten mit ordinalskalierten Variablen gilt dann als akzeptabel, wenn angenommen werden kann, dass die ordinale Messung lediglich eine schlechte Messung einer 'wahren' stetigen Skala darstellt. Dies ist normalerweise beispielsweise bei Einstellungs- oder Meinungs-Items der Fall, von denen angenommen wird, dass die befragte Person sich mit ihrer Einstellung auf einer stetigen Skala befindet, aber aufgrund der vorgegebenen Skala (etwa von eins bis fünf) lediglich wenige Punkte auf dieser Skala angeben darf (University of Texas, w/o year). In ähnlicher Weise lässt sich das Geschlecht als schlechte Messung der stetigen Variable 'soziales Geschlecht' (Gender) verstehen. Hier ist allerdings die Diskrepanz zwischen der 'wahren' Variable Gender und der gemessenen Variable Geschlecht besonders groß, da das Geschlecht lediglich zwei Ausprägungen hat.

Das Argument der ungleichen Abstände zwischen den Ausprägungen ordinaler Variablen ist zwar formal korrekt. Allerdings lässt sich dem zumindest bei Befragungs-Items entgegenhalten, warum eine befragte Person die Abstände zwischen den Ausprägungen deutlich unterschiedlich interpretieren sollte. Warum sollte ein Befragter etwa die Unterschiede zwischen "sehr wichtig" und "wichtig" sowie zwischen "wichtig" und "weniger wichtig" als unterschiedlich groß empfinden? In den StadtLeben-Daten wurden die Formulierungen der Ausprägungen so gewählt, dass man nach experimentellen Methodenstudien (Rohrmann 1978) von gleichen Abständen ausgehen kann.

Dies gilt allerdings nicht für die Pkw-Verfügbarkeit. Aber auch hier dürfen cum grano salis die Abstände

zwischen den Ausprägungen als gleich angesehen werden, wenn man die Abstände zwischen den tatsächlichen Nutzungsgewohnheiten des Pkw zugrunde legt (s.o.).

Zusammenfassend lässt sich festhalten, dass das Skalierungsniveau aller verwendeten Variablen der Anwendung der ML-Schätzmethode nicht zwingend entgegensteht.

Als weiteres Problem ist die Schiefe einiger Verteilungen zu erwähnen. Dies betrifft insbesondere die zurückgelegten Distanzen, bei den Kurtosis-Werte  $>10$  auftreten. Diese lassen sich durch Logarithmierung deutlich reduzieren, wie dies in der Verkehrsforschung häufig gemacht wird.

## 6.3 Angewandtes Schätzverfahren und Validierung der Ergebnisse

Nach den oben diskutierten Ergebnissen ist die vorhandene Stichprobengröße von rund  $n=2.000$  bei Weitem ausreichend, um das ML-Schätzverfahren als robust ansehen zu können. Dies gilt auch, wenn die Stichprobe in eine Hauptstichprobe und eine Validierungsstichprobe geteilt wird. Das ADF-Verfahren gelangt dann an die Grenze seiner Verlässlichkeit, ist aber ebenfalls noch in den meisten Fällen akzeptabel. So ist bei einigen Modellen die Anzahl der Freiheitsgrade bei ca.  $df=50$  bis  $df=75$ , so dass die 'Daumenregel'  $n > 20 * df$  bei  $n=1.000$  gerade noch erfüllt ist oder nur knapp verfehlt wird.

Die Schiefe der Verteilung einiger Variablen spricht zunächst für die Anwendung des ADF-Verfahrens. Nach den oben berichteten Ergebnissen von Hoogland und Boomsma (1998) lässt sich allerdings aufgrund der großen Stichprobe auch das ML-Verfahren ohne größere Verzerrungen anwenden.

In der Konsequenz kommen zur besseren Absicherung der Ergebnisse beide Verfahren zum Einsatz. Sie sind Bestandteil eines recht rigorosen Ansatzes zur Validierung der Ergebnisse, für den drei Gründe ausschlaggebend waren:

- der erwähnte Mix aus theorie- und datengeleiteter Modellierung
- die Sensibilität der Modellschätzungen gegenüber Modellvariationen
- es sind nicht alle statistischen Voraussetzungen für die verwendete Methodik erfüllt.

Als Konsequenz wurde zunächst die Stichprobe zufällig in zwei Hälften geteilt. Dann wurde jedes Modell in vier Versionen geschätzt:

1. ML-Schätzung des theoretischen Modells mit der Hauptstichprobe
2. Empirische Anpassung des Modells an die Daten

3. ADF-Schätzung des theoretischen Modells
4. ML-Schätzung des theoretischen Modells mit der Validierungsstichprobe

Die Version 2 dient lediglich der Verifizierung der Schätzung von Koeffizienten im theoretischen Modell, wenn dieses an die Daten angepasst wird. Die Modellanpassung erfolgt durch schrittweises Freisetzen von Fehlerkovarianzen auf der Grundlage von Modifikationsindizes. Die Modifikationsindizes geben die erwartete Verbesserung des Chi-Quadrat-Wertes an, wenn ein bestimmter Pfad im Modell freigesetzt wird. Die Modellanpassung wird abgebrochen, wenn eine akzeptable Modellanpassung erreicht ist, weil das inhaltliche Hauptinteresse nicht in einer maximalen Modellanpassung, sondern in der Stabilität und theoretischen Plausibilität der theoretischen Modelle liegt.

Allerdings ist auch in der Spezifikation der theoretischen Modelle die Vorgehensweise nach der "reinen Lehre" nicht immer möglich, weil bei substanziell gleichen Modellen, aber unterschiedlichen Schätzmethoden oder Stichproben gelegentlich Ergebnisse auftreten, die einen direkten Vergleich nicht zulassen. Diese resultieren insbesondere aus dem Messmodell der Lebenslage. So entsteht hin und wieder empirische Unteridentifizierung, d.h. ein Modell ist theoretisch, aber nicht mit den gegebenen Daten lösbar. Dies führt dazu, dass beispielsweise ein und das gleiche Modell mit der ersten Stichprobe lösbar ist, jedoch nicht mit der Validierungsstichprobe. Deshalb ist es auch in den theoretisch hergeleiteten Modellen hin und wieder notwendig, in sehr sparsamem Ausmaß Kovarianzen zwischen Fehlervariablen der Lebenslage freizusetzen. Insofern sind die theoretischen Modelle nicht vollständig deduktiv entwickelt, sondern enthalten einige 'empiristische' Elemente.

Die vier Versionen eines Modells wurden im Hinblick auf Stärke und Richtung der auftretenden Effekte miteinander verglichen. Da viele Beziehungen zwischen je zwei Variablen in identischer Form in einer großen Anzahl an Modellen untersucht wurde (z.B. die Effekte der Lebenslage auf Wohnortpräferenzen), wurde das Gesamtbild der Effekte auch benutzt um einen generellen Eindruck von der Stabilität der Modelle zu gewinnen. Im Ergebnis zeigen sich durchaus starke Variationen zwischen je vier Versionen eines Modells im Sinne der direkten Effekte einer Variablen auf eine andere Variable. Werden allerdings die Gesamteffekte berücksichtigt (einschließlich indirekter Effekte), dann erweisen sich die Ergebnisse als recht stabil und sind im Sinne von Effektrichtungen und -stärken gut interpretierbar.

## 6.4 Ausblick

Bis zum Zeitpunkt der Erstellung dieses Papiers liegen aus den Analysen ohne latente Variablen bereits mehrere Veröffentlichungen vor (Scheiner 2006b,

Scheiner 2007a, Scheiner/Holz-Rau 2007). Aus den Analysen mit latenten Variablen sind Veröffentlichungen in Vorbereitung bzw. wurden bereits auf Tagungen vorgestellt (Scheiner 2007c, Scheiner 2007d, Scheiner 2007e).

Weiterentwicklungen des hier vorgestellten Ansatzes könnten sich auf drei Schwerpunkte konzentrieren:

- Systematische Integration weiterer Bausteine in die Modelle, etwa Wohn- und Standortzufriedenheit sowie Einstellungen zu Verkehrsmitteln (Mobilitätsstile). Letztere könnten ähnlich wie Standortpräferenzen als intervenierende Größen in die Modelle eingebaut werden. Dabei dürfte allerdings das Risiko von Tautologien zwischen der Messung von Einstellungen und Handeln ungleich größer sein als im Falle von Standortpräferenzen und Standortverhalten. Dies ist darin begründet, dass Verkehrsmittel im hiesigen gesellschaftlichen Kontext mehr oder weniger ubiquitär verfügbar sind und eine Person mit einer positiven Einstellung beispielsweise zum Pkw mit hoher Wahrscheinlichkeit diese Präferenz ohne Weiteres realisieren kann. Bei Standortpräferenzen gilt dies aufgrund der Restriktionen des Wohnungsmarktes keineswegs; deshalb ist eine vergleichbar enge Beziehung zwischen Standorteinstellung und tatsächlich getroffener Standortwahl nicht unbedingt zu erwarten. Bei der Integration von Maßen der *Zufriedenheit* steht ein großes Fragezeichen hinter der Kausalität. Ben-Akiva (2007) sieht beispielsweise die Zufriedenheit mit Aktivitäten als ursächlich für die Generierung von außerhäuslichen Aktivitäten und damit Wegen an. Ein hohes Maß außerhäuslicher Aktivitäten kann aber auch Zufriedenheit mit diesen Aktivitäten *hervorrufen*.
- Systematische Untersuchungen alternativer Kausalstrukturen zwischen den Bausteinen, etwa zwischen Pkw-Verfügbarkeit, Standortpräferenzen und tatsächlich getroffenen Standortentscheidungen. Sollte sich beispielsweise herausstellen, dass Standortverlagerungen eine Veränderung von Standortpräferenzen nach sich ziehen, dann spräche das für eine stärkere Bedeutung objektiver Raumstrukturen gegenüber Standortpräferenzen und für eine geringere Bedeutung residenteller Selbstselektion.
- Neues Licht auf Kausalstrukturen könnten auch Panelmodelle werfen, in denen die wechselseitigen Beziehungen zwischen Modellbausteinen im Zeitverlauf besser abgebildet werden können als in querschnittsorientierten Modellen. Panelmodelle könnten darüber hinaus zu einem besseren Verständnis von Mobilitätsbiographien beitragen, die in den letzten Jahren verstärkt diskutiert werden (Lanzendorf 2003, Scheiner 2007b).

## 7 Literatur

- Ajzen, Icek (1991). The Theory of Planned Behavior. In: *Organizational Behavior and Human Decision Processes* 50, S. 179-211.
- Bamberg, Sebastian (2004): Sozialpsychologische Handlungstheorien in der Mobilitätsforschung. Neuere theoretische Entwicklungen und praktische Konsequenzen. In: Dalkmann, Holger / Lanzendorf, Martin / Scheiner, Joachim (Hg.): *Verkehrsgenese: Entstehung von Verkehr sowie Potenziale und Grenzen der Gestaltung einer nachhaltigen Mobilität. Studien zur Mobilitäts- und Verkehrsforschung* 5. Mannheim. S. 51-70.
- Beckmann, Klaus J. / Hesse, Markus / Holz-Rau, Christian / Hunecke, Marcel (Hg., 2006): *StadtLeben – Wohnen, Mobilität und Lebensstil. Neue Perspektiven für Raum- und Verkehrsentwicklung*. Wiesbaden.
- Ben-Akiva, Moshe (2007): Happiness, Activity-Based Transport Modelling. Keynote lecture at the European Transport Conference, Leiden, NL, 17-19 October 2007.
- Bentler, Peter M. / Chou, Chih-Ping (1987): Practical Issues in Structural Modeling. In: *Sociological Methods & Research* 16(1), S. 78-117.
- Berger, Peter A. / Hradil, Stefan (1990): Die Modernisierung sozialer Ungleichheit – und die neuen Konturen ihrer Erforschung. In: Berger, Peter A. / Hradil, Stefan (Hg.): *Lebenslagen, Lebensläufe, Lebensstile. Soziale Welt, Sonderband 7*. Göttingen. S. 3-24.
- Bollen, Kenneth A. (1989): *Structural Equations with Latent Variables*. New York: Wiley.
- Bram, Jason / McKay, Alisdair (2005): The Evolution of Commuting Patterns in the New York City Metro Area. In: *Current Issues in Economics and Finance* 11(10), S. 1-7.
- Browne, Michael W. (1984): Asymptotic Distribution Free Methods in the Analysis of Covariance Structures. In: *British Journal of Mathematical and Statistical Psychology* 37, S. 62-83.
- Bureau van Dijk Electronic Publishing/Creditreform (2004): *Markus. Marketinguntersuchungen (CD-Rom)*. Neuss.
- Cao, Xinyu / Mokhtarian, Patricia L. / Handy, Susan L. (2007): Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. In: *Transportation* 34(5), S. 535-556.
- Golob, Thomas F. (2001). *Structural Equation Modeling for Travel Behavior Research*. Working Paper UCI-ITS-AS-WP-01-2. Irvine, CA: Institute of Transportation Studies, University of California.
- Hammer, Antje / Scheiner, Joachim (2006): *Lebensstile, Wohnmilieus, Raum und Mobilität*. In: Beckmann/Hesse/Holz-Rau/Hunecke 2006, S. 15-30.
- Holz-Rau, Hans-Christian (1991): Verkehrsverhalten beim Einkauf. In: *Internationales Verkehrswesen* 43(7-8), S. 300-305.
- Hoogland, Jeffrey J. / Boomsma, Anne (1998) *Robustness Studies in Covariance Structure Modeling: An Overview and a Meta-Analysis*. In: *Sociological Methods & Research* 26(3), S. 329-367.
- Hunecke, Marcel / Schweer, Indra (2006): Einflussfaktoren der Alltagsmobilität – Das Zusammenwirken von Raum, Verkehrsinfrastruktur, Lebensstil und Mobilitätseinstellungen. In: Beckmann/Hesse/Holz-Rau/Hunecke 2006, S. 147-166.
- Johnson, Rucker C. (2003): *Landing a Job in Urban Space: The Extent and Effects of Spatial Mismatch*. Unpublished manuscript. [www.fordschool.umich.edu/research/poverty/pdf/SMHRSUE.pdf](http://www.fordschool.umich.edu/research/poverty/pdf/SMHRSUE.pdf) (15.2.2006)
- Kasper, Birgit / Scheiner, Joachim (2006): Räumliche Mobilität als Prozess kurz- und langfristigen Handelns: Zusammenhänge zwischen Wohn- und Alltagsmobilität. In: Beckmann/Hesse/Holz-Rau/Hunecke 2006, S. 167-186.
- Kühnel, Steffen (2006): *Einführung in die Analyse linearer Strukturgleichungsmodelle mit LISREL 8.80 (Materialien zum Workshop)*. Mannheim.
- Lanzendorf, Martin (2003): *Mobility Biographies. A New Perspective for Understanding Travel Behaviour*. Paper presented at the 10th International Conference on Travel Behaviour Research (IATBR), Lucerne, 10th-15th August 2003.
- MacCallum, Robert C. / Browne, Michael W. / Sugawara, Hazuki M. (1996): Power Analysis and Determination of Sample Size for Covariance Structure Modeling. In: *Psychological Methods* 1(2), S. 130-149.
- Müller, Hans-Peter (1992): *Sozialstruktur und Lebensstile*. Frankfurt / Main.
- Rohrmann, Bernd (1978): Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. In: *Zeitschrift für Sozialpsychologie* 9(1), S. 222-245.
- Scheiner, Joachim (2005a): Auswirkungen der Stadt- und Umlandwanderung auf Motorisierung und Verkehrsmittelnutzung: ein dynamisches Modell des Verkehrsverhaltens. In: *Verkehrsforschung Online* 1(1), S. 1-17.
- Scheiner, Joachim (2005b): *Methodische Anmerkungen zu den Analysen der Haushaltsbefragung im Projekt StadtLeben. Raum und Mobilität – Arbeitspapiere des Fachgebiets Verkehrswesen und Verkehrsplanung* 11. Dortmund.

- Scheiner, Joachim (2006a): Standortbewertungen und Wohnmobilität. In: Beckmann/Hesse/Holz-Rau/Hunecke 2006, S. 67-78.
- Scheiner, Joachim (2006b): Wohnen und Aktionsraum: Welche Rolle spielen Lebensstil, Lebenslage und Raumstruktur? In: Geographische Zeitschrift 94(1), S. 43-62.
- Scheiner, Joachim (2007a): Einkaufsverkehr im räumlichen und sozialen Kontext: Die Bedeutung von Wohnstandortwahl, Lebenslage und Lebensstil. Erscheint in: Studien zur Mobilitäts- und Verkehrsforschung (zur Veröffentlichung angenommen)
- Scheiner, Joachim (2007b): Mobility Biographies: Elements of a Biographical Theory of Travel Demand. In: Erdkunde 61(2), S. 161-173.
- Scheiner, Joachim (2007c): Travel behaviour - evidence for objective and subjective inequalities. Paper presented at the Cosmopolitanities network meeting, Basle, CH, 7-8 September 2007.
- Scheiner, Joachim (2007d): Far, far away - micro-spatial analyses of trip distances and mode choice. Paper presented at the European Transport Conference, Leiden, NL, 17-19 October 2007.
- Scheiner, Joachim (2007e): Travel behaviour – affected by objective spatial context and/or by subjective accessibility preferences?" Paper presented at the European Science Foundation (ESF) Workshop "How to Measure Access", Dresden, 28 September 2007.
- Scheiner, Joachim / Holz-Rau, Christian (2007): Travel Mode Choice: Affected by Objective or Subjective Determinants? In: Transportation 34(4), S. 487-511.
- Scheiner, Joachim / Kasper, Birgit (2005): A Lifestyles Approach to Investigating Residential Mobility and Travel Behaviour. In: Williams, Katie (Hg.): Spatial Planning, Urban Form and Sustainable Transport. Aldershot. S. 42-60.
- Schermelleh-Engel, Karin / Moosbrugger, Helfried / Müller, Hans (2003): Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. In: Methods of Psychological Research Online 8, S. 23-74.
- Schneider, Nicole / Spellerberg, Annette (1999): Lebensstile, Wohnbedürfnisse und räumliche Mobilität. Opladen.
- Schweer, Indra R. / Hunecke, Marcel (2006): Die Lebensstile in StadtLeben. In: Beckmann/Hesse/Holz-Rau/Hunecke 2006, S. 55-61.
- Simma, Anja (2000): Verkehrsverhalten als eine Funktion soziodemografischer und räumlicher Faktoren. Arbeitsbericht Verkehrs- und Raumplanung 55 des Instituts für Verkehrsplanung, Transporttechnik, Straßen- und Eisenbahnbau der ETH Zürich. Zürich.
- Simma, Anja / Axhausen, Kay W. (2001): Structures of Commitment in Mode Use: A Comparison of Switzerland, Germany and Great Britain. In: Transport Policy 8(4), S. 279-288.
- Spellerberg, Annette (1996): Soziale Differenzierung durch Lebensstile. Eine empirische Untersuchung zur Lebensqualität in West- und Ostdeutschland. Berlin.
- Stadt Köln, Der Oberbürgermeister (2005): Weiterführende Schulen in Köln. Sekundarstufe I. Köln.
- Stevens, James P. (1996): Applied Multivariate Statistics for Social Sciences. New Jersey.
- Titheridge, Helena (2000): The People: Where will they travel (and how)? In: Banister, David / Stead, Dominic (Hg.): Evidence from the University College London, Bartlett School of Planning to the Environmental Planning Study of the Royal Commission on Environmental Pollution. Annex 2. <http://www.rcep.org.uk/epevid/p2-ucl.htm> (15.2.2006)
- University of Texas (w/o year): Structural Equation Modeling Using AMOS: An Introduction. ([www.utexas.edu](http://www.utexas.edu), 25.1.2007)
- Wheaton, Blair (1988): Assessment of Fit in Overidentified Models with Latent Variables. In: Long, J. Scott (Ed.): Common Problems / Proper Solutions: Avoiding Error in Quantitative Research. Newbury Park, CA: Sage. S. 193-225.
- Wieland, Dirk (2004): Die Grenzen der Individualisierung. Opladen.
- Zängler, Thomas W. (2000): Mikroanalyse des Mobilitätsverhaltens in Alltag und Freizeit. Berlin.