

Stefan Schönfelder und Kay W. Axhausen

## Periodizität im Verkehrsverhalten: Erste Ergebnisse mit Überlebenszeitmodellen

*Bei den Forschungen zu zeitlichen Aspekten des Verkehrsverhaltens auf der Ebene von Personen und Haushalten haben Untersuchungen zu rhythmischen Mustern der Mobilität bisher eine untergeordnete Rolle gespielt. Mit dem Forschungsprojekt Mobidrive steht erstmals ein Datensatz zum langfristigen Raum-Zeit-Verhalten von Verkehrsteilnehmerinnen und -teilnehmer zur Verfügung, der es erlaubt die Periodizität auf der Wochen und Monatsebene abzubilden. In diesem Aufsatz wird ein vielversprechender Ansatz zur Charakterisierung und Erklärung von rhythmischen Mustern vorgestellt, der auf einer Anwendung von stochastischen Modellen aus der Überlebenszeitanalyse (Survival Analysis) beruht.*

**Schlagwörter:** Rhythmische Muster, Periodizität, Verkehrsverhalten, Survival Analysis, Modell-schätzung

### 1. Rhythmische Muster des Verkehrsverhaltens

Zeitliche Aspekte des Verkehrsverhaltens von Personen und Haushalten bilden seit den Anfängen des *aktivitäten-basierten Ansatzes* der Verkehrsforschung einen wesentlichen Schwerpunkt dieser Forschungsrichtung (vgl. Jones, 1981; Kitamura, 1988; Mahmassani, 1988; Pas und Harvey, 1997). Dabei wurden bisher insbesondere der Charakter sowie die Bestimmungsgrößen von

- Verhaltensroutinen,
- Verhaltensvariabilität und
- Verhaltensdynamik

eingehend untersucht. Konzeptionelle Grundlage des *aktivitäten-basierten Ansatzes* ist die Vorstellung, dass individuelle Ortsveränderungen das Bedürfnis nach ausserhäuslichen Aktivitäten befriedigen. Verkehr wird somit induziert und kann als abgeleitete Grösse angesehen werden. Ortsveränderungen unterliegen diversen individuellen Restriktionen des Verkehrsteilnehmers sowie den räumlichen, aber vor allem zeitlichen Strukturen und Vereinbarungen der dynamischen Umwelt (vgl. Hägerstrand, 1970).

Mit den rhythmischen Mustern des Verkehrsverhaltens soll im folgenden eine weitere zeitliche Perspektive im Vordergrund stehen, die in den bisherigen Analysen des *aktivitäten-basierten Ansatzes* erst wenig Beachtung gefunden hat. Für verkehrsrelevante Planungen und Massnahmen öffentlicher sowie privater Akteure ist das Verständnis individu-

eller Mobilitätsgewohnheiten wie rhythmischer Muster eine essentielle Entscheidungsgrundlage. Beispielsweise zielt der Forschungsbereich *Mobilität und Verkehr besser verstehen* des Bundesministeriums für Bildung und Forschung mit einer Reihe von Projekten darauf ab, einerseits zur Weiterentwicklung des methodischen Analyseinstrumentariums beizutragen und andererseits die Einflussgrößen der alltäglichen Mobilität vertieft zu untersuchen bzw. bisher nicht berücksichtigte zu eruieren<sup>1</sup>.

Rhythmische Muster werden als wiedererkennbare Elemente verstanden, die in regelmässigen Abständen im Verkehrsverhalten von Personen und Haushalten beobachtet werden können. In diesem Aufsatz werden insbesondere Rhythmen angesprochen, die über längere Zeiträume andauern (Wochen- oder Monatsebene) und damit weit über den üblichen Beobachtungszeitraum eines Stichtages hinausgehen. Zu den rhythmischen Mustern gehören komplette Tagesprogramme mit übereinstimmenden Attributen oder aber Teilsequenzen bzw. einzelne Hauptaktivitäten oder -wege. Die wiedererkennbaren Elemente erscheinen meist in periodischer, voraussagbarer Weise und können historisch erklärt werden, d.h. sie sind das Resultat koordinierter Rhythmen und Strukturen der menschlichen Physiologie sowie der dynamischen Umwelt (Shapcott und Steadman, 1978).

<sup>1</sup> Siehe dazu <http://www.tuev-ptmuv.com/tf5.htm>

### Mobidrive – Dynamik und Routinen im Verkehrsverhalten

Der Identifizierung und der Analyse von rhythmischen Mustern des Verkehrsverhaltens stand bisher die unzureichende Verfügbarkeit von Langzeitbeobachtungen entgegen. Die Erhebungszeiträume üblicher Wegetagebuchehebungen gehen in der Regel nicht über eine Woche hinaus –beschränken sich meist sogar auf nur einen Stichtag (z.B. KON-TIV). Das Datenmaterial solcher Erhebungen liess bislang keine sinnvollen Untersuchungen von Verhaltensrhythmen zu, deren Periodizität sich auf der Wochen- oder gar Monatsebene bewegt. Die einzig bekannte Wegetagebuchehebung, die einen mehrwöchigen, zusammenhängenden Erhebungszeitraum aufweist, ist der *Uppsala Household Travel Survey* von 1971 (Hanson und Burnett, 1981; Hanson und Hanson, 1981a, 1981b). Dessen Daten entsprechen aufgrund ihres Alters und der lückenhaften Verfügbarkeit nur eingeschränkt den Ansprüchen aktueller Forschungsfragen.

Mit der Durchführung einer sechswöchigen Haushalts- und Personenbefragung zum Verkehrsverhalten im Forschungsprojekt *Mobidrive* – Dynamik und Routinen im Verkehrsverhalten – steht nunmehr ein aktueller deutscher Datensatz zum langfristigen Verkehrsverhalten zur Verfügung (vgl. Axhausen, Zimmermann, Schönfelder, Rindsfuser und Haupt, 2000). Die Erhebung wurde von den Projektpartnern im Herbst 1999 in den Städten Halle/Saale und Karlsruhe durchgeführt und wurde um Rahmendaten zur Sozio-Demographie sowie um zusätzliche Informationen zu mobilitätsspezifischen

und allgemeinen Einstellungen der befragten Haushalte ergänzt.

Die im folgenden dargelegte Entwicklung eines stochastischen Modells zur Abbildung rhythmischer Muster basiert auf dem gesammelten Datenmaterial und ist im direkten Zusammenhang mit dem Forschungsprojekt *Mobidrive* entstanden. *Mobidrive* wird im Rahmen der oben genannten Mobilitätsforschungsinitiative des Bundesministeriums für Bildung und Forschung gefördert und will einen Beitrag zur intensiven Untersuchung und zur Erklärung von Prozessen leisten, die die Bildung von Routinen und Veränderungen im Verkehrsverhalten beeinflussen.

Die ersten deskriptiven Untersuchungen des deutschland- und weltweit einmaligen Datensatzes belegen den weitgehend routinisierten Charakter des Verkehrsverhaltens und den Einfluss fester zeitlicher Strukturen (z.B. Arbeitszeitregime oder Ferienzeiten) innerhalb eines Tages und den Wo-

chenverlauf. Weiterhin ergeben sich interessante Unterschiede bei den Zeitznutzungsstrukturen zwischen Ost- und Westdeutschland (Abbildung 1).

## 2. Auswahl und Grundlagen eines geeigneten Analyseinstruments: Überlebenszeitmodelle (*Survival Analysis*)

Eine über die deskriptive Beschreibung hinausgehende Analyse des Charakters von rhythmischen Mustern des Verkehrsverhaltens und eine Identifizierung von Determinanten stellt die Verkehrsforschung vor methodische Herausforderungen. Auf der aggregierten Ebene der Zähldaten (Systemebene) wird seit langem mit Zeitreihenmodellen gearbeitet, bei denen zeitlich geordnete Folgen von Beobachtungen (z.B. Verkehrsstärken an Zählstellen) einer Größe beschrieben, modelliert, prognostiziert und auf deren Konsistenz kontrolliert werden (vgl. Schlittgen und Streitberg, 1999). Für den aktivitäten-basierten Ansatz mit individuell-zuordbaren Werten und Kennziffern ist jedoch eine methodische Weiterentwicklung erforderlich.

Rhythmen können generell durch ihre Perioden und Amplituden typisiert werden (Abbildung 2). Dieser Aufsatz beschränkt sich auf die Analyse der Periodizität des Verkehrsverhaltens, die durch die Anordnung und Länge der Zeitintervalle zwischen gleichartigen Mustern (hier: Aktivitäten) be-

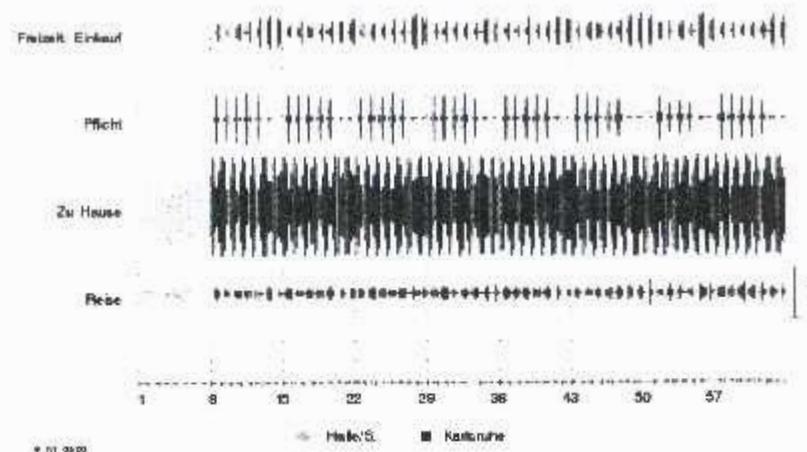


Abb. 1: Zeitnutzung auf Sechs-Wochen-Ebene (aggregiert)

schrieben werden kann. Die Ausprägungen der rhythmischen Muster (Amplituden) können beliebig komplex gewählt werden – von der eher eindimensionalen Betrachtung der Aktivitäten mit lediglich einem Attribut (z.B. Zweck) bis zu einer komplexen multiplen Charakterisierung von Mustern nach verschiedenen Attributen. Frühere Forschungsarbeiten stellen eine Reihe Mustervergleichs- bzw. Musterklassifizierungsmethoden für Wege, Aktivitäten oder Sequenzen zur Verfügung (z.B. Pas, 1983; Recker, McNally und Root, 1983), die in eine vertiefte Analyse einfließen könnten.

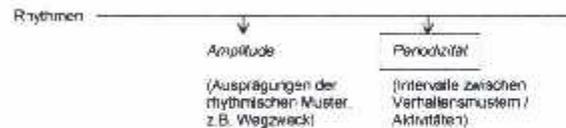


Abb. 2: Strukturelle Elemente von Rhythmen

Für die Analyse von Intervallen zwischen gleichartigen Verhaltensmustern wurden aus dem Mobidrive-Datenmaterial die Zeitdauern zwischen Aktivitäten gleicher Art – kategorisiert nach Aktivitätszwecken – generiert (Abbildung 3). In einem weiteren Schritt wurden diese Daten mit soziodemographischen Charakteristika der Reisenden

No.	H	P	Activity	Day	Activity	Day	Interval	Interval
	H		Start Time		End Time		[hours]	[days]
1	3	1	15:35:00	1	17:00:00	1		
2	3	1	08:30:00	3	09:15:00	3	40.917	1.7049
3	3	1	15:25:00	8	16:30:00	8	126.917	5.2882
4	3	1	09:35:00	12	10:30:00	12	90.167	3.7569
1	2	2	15:35:00	1	17:00:00	1		
2	3	2	15:25:00	8	16:30:00	8	167.833	6.9931
3	3	2	15:30:00	9	16:30:00	9	24.083	1.0035
4	3	2	12:20:00	10	13:30:00	10	20.833	0.8681
5	3	2	09:35:00	12	11:00:00	12	45.750	1.8854

\* - Beobachtete Beobachtungen, siehe unten; H1 = Hausnummer, P = Personnummer

Abb. 3: Intervalle zwischen gleichartigen Aktivitäten, Aktivität: Langfristiger Einkauf (Ausschnitt eines SAS-Output)

und weiteren Attributen der Aktivitäten verknüpft, die als unabhängige Variablen in das vorgeschlagene Modell eingehen (siehe Abschnitt 4).

**Besonderheiten bei der Analyse von Zeitdauerdaten**

Lineare Regressionsmodelle und verwandte stochastische Werkzeuge bieten sich in der Regel zur Untersuchung von Verteilungen und den Einfluss bzw. die Korrelation von Determinanten eines bestimmten Phänomens an. Bei der Analyse von Dauerdaten ist die Anwendung solcher Modelle mit Kleinst-Quadrato-Schätzung aufgrund von Annahmen zu den Modellergebnissen und aufgrund des besonderen Charakters der empirischen Daten jedoch problematisch (Hosmer und Lemeshow, 1999). Empirische Dauerdaten und somit auch Ergebnisse von Schätzungen können ausschliesslich positive Werte annehmen – negative Zeitdauern existieren nicht! Diese Annahme stellt eine wesentliche Restriktion für den Einsatz klassischer Regressionen dar.

Zudem können Zeitdauern aus vielerlei Gründen häufig nur teilweise beobachtet und gemessen werden – eine Tatsache, die in der Literatur als *censoring problem* bekannt ist und die die Anwendung klassischer Regressionsmodelle ausschliesst. Die Dauer des Beobachtungszeitraums schränkt in der Regel die Plausibilität der Messungen ein, da die relevanten Prozesse zum Teil schon vor dem Beginn der Untersuchungen eingesetzt haben oder über deren Ende hinausgehen können. Im ersten Fall spricht man von *left censoring*, d.h. dass der wahre Beginn einer Dauer ignoriert und dagegen der Start des Beobachtungszeitraums als Anfangs-

zeitpunkt definiert (*zensiert*) wird. Über die beobachtete Zeitdauer lässt sich lediglich aussagen, dass sie kleiner als der wahre Wert ist. *Right censoring* beschreibt das Gegenteil, nämlich die Tatsache, dass der in Wirklichkeit abgelaufene Prozess länger ist als der beobachtete bzw. definierte Prozess innerhalb der Untersuchung (vgl. dazu Allison, 1995, 9ff.). Beispiel: Der zeitlich nicht präzisierbare Tod eines Patienten nach dem Ende des Untersuchungszeitraums wird ignoriert und plausibel festgelegt. Herkömmliche statistische und stochastische Methoden erübrigen sich somit für ernstzunehmende Analysen, da beispielsweise eine Mittelwertberechnung der Dauern quasi unmöglich ist oder Variablenausprägungen wie „mindestens 5 Monate Überlebenszeit“ nicht in stochastische Modelle einfließen können.

Diese statistischen und stochastischen Restriktionen haben zur Entwicklung der sogenannten Überlebenszeitanalyse oder *Survival Analysis* geführt, mit deren Hilfe hier auch ein geeignetes Modell für die in Mobidrive generierten Intervalldaten entwickelt werden soll<sup>2</sup> (Kalbfleisch und Prentice, 1980; Cox, 1984; Kleinbaum, 1996). Das Ziel von *Survival Analysis* ist die beobachtete Verteilung von Dauern mit Hilfe von sogenannten Hazard-Funktionen abzubilden (siehe unten).

*Survival analysis* findet in anderen technischen und gesellschaftlichen Forschungsfeldern wie der Biometrie, dem Maschinenbau oder der Marktforschung schon länger Anwendung. *Survival Analysis*

<sup>2</sup> Im Folgenden wird durchgehend der englische Begriff *Survival Analysis* benutzt.

und insbesondere sogenannte *hazard models* werden überall dort eingesetzt, wo der Zeitpunkt des Eintritts eines bestimmten Ereignisses prognostiziert werden soll. Als typische Anwendungsfälle gelten die Prognose des Andauerns von Arbeitslosigkeit bis zur Wiederaufnahme einer Beschäftigung, des Zeitraums bis zur Marktdurchdringung eines neuen Produktes oder der Länge eines Arbeitskampfes bis zu dessen Schlichtung. Mit *hazard models* wird in der Verkehrsverhaltensforschung seit etwa Ende der 80er Jahre gearbeitet, z.B. in den Bereichen Aktivitätenwahlverhalten oder der Analyse von Dauern in- und aushäusiger Aktivitäten (Mannering und Hamed, 1990; Hamed und Mannering, 1993; Mannering, Murakami und Kim, 1992; Hamed, Kim und Mannering, 1992; Niemeier und Morita, 1994; Bhat, 1996a; 1996b; Ettema, Borgers und Timmermans, 1995; Reader und McNeill, 1999; Oh, 1999).

#### Konzeptionelle Grundannahmen zur Periodizität des Verkehrsverhaltens

*Survival Analysis* stellt unseres Erachtens einen lohnenden Ansatz zur Analyse der Periodizität im Verkehrsverhalten dar. Konzeptionelle Grundlage des hier vorgestellten Ansatzes ist die Vorstellung, dass rhythmische Muster als Ereignisse angesehen werden können, deren Eintrittswahrscheinlichkeiten über einen Zeitraum annähernd genau bestimmt werden können. In allererster Linie werden die Wahrscheinlichkeiten über die zeitlichen Abstände (Intervalle) zum letztmaligen gleichartigen Muster bestimmt. Weitere Voraussetzung für die Schätzung der Ereigniszeitpunkte ist die Identifizierung von weiteren Determinanten des Verhaltens, die Einfluss

auf den zyklischen Charakter der Aktivität ausüben besitzen. Im (Zeit-)Intervall zwischen zwei Ereignissen findet ein Bedürfnisaufbau statt, der durch eine Wahrscheinlichkeitsfunktion ausgedrückt werden kann und an dessen Ende (relatives Maximum) der Ereigniseintritt steht.

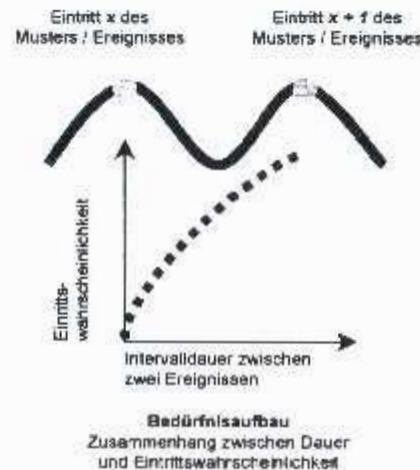


Abb. 4: Basiskonzept: Rhythmik des Verkehrsverhaltens

#### Mathematische Grundlagen der *Survival Analysis*

Innerhalb der *Survival Analysis* wurden eine Reihe von Modellen entwickelt, die sich durch die Verteilungsannahmen zur individuellen Ereigniszeit  $T$  unterscheiden. Die Zeit (Zeitpunkt)  $T$  ist auch die im Mittelpunkt stehende Zufallsvariable unseres Modells – sie beschreibt den Zeitpunkt des Beginns

einer gleichartigen Aktivität bzw. den Endpunkt eines Intervalls zwischen gleichartigen Mustern.

Für das bessere Verständnis des im folgenden Abschnitt beschriebenen stochastischen Modells, sollen an dieser Stelle drei wesentliche Darstellungsformen der Wahrscheinlichkeitsverteilungen von  $T$  vorgestellt werden (Abbildung 5):

Zunächst liefert uns die (kumulative) Verteilungsfunktion  $F(t)$  die Wahrscheinlichkeit, dass der Zeitpunkt des Ereigniseintritts vor oder genau an einem gewählten Zeitpunkt  $t$  liegt. Könnten wir den Wert von  $F$  für jeden Wert von  $t$  bestimmen, wüssten wir alles wesentliche über die Verteilung von  $T$  und damit über die allgemeine Verteilung der Intervall-dauern der entsprechenden Rhythmen.

$$F(t) = \Pr[T < t] \quad (1)$$

Üblicherweise steht in Anwendungen von *Survival Analysis* das Überleben von Prozessen über den gewählten Zeitpunkt – also beispielsweise dem Ende des Untersuchungszeitraum – im Mittelpunkt des Interesses. Die Survival Funktion  $S(t)$  drückt die Wahrscheinlichkeit des Überlebens über einen Zeitpunkt  $t$  hinaus aus. Die Funktion kann Werte zwischen 0 und 1 annehmen, und da Zeitdauern per Definition nicht-negativ sind, ergibt sich  $S(0) = 1$ . Die Form hat je nach Eigenschaft der beobachteten Prozesse verschiedenste Ausprägungen, ist aufgrund der beschriebenen Restriktionen in der Regel jedoch fallend.

$$S(t) = \Pr[T \geq t] = 1 - F(t) \quad (2)$$

In engem Zusammenhang mit (1) steht die unbestimmte Dichtefunktion bzw. *Probability Density*

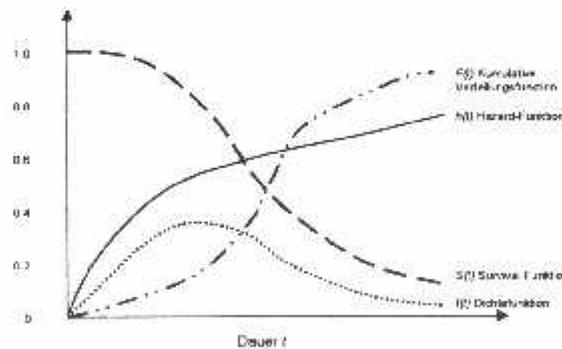
*Function.* Als erste Ableitung der Kumulativen Verteilungsfunktion entspricht sie weniger der intuitivem Vorstellung von Verteilungsausprägungen oder – kurven. Sie spielt jedoch bei der späteren Modellentwicklung eine wichtige Rolle, da Annahmen zur unbestimmten Verteilung von T wesentlichen Einfluss auf die auf die Schätzwerte für die Intervall-dauern haben. Sie ist gegeben durch

$$f(t) = dF(t) / d(t) = - dS(t) / dt = \lim_{dt \rightarrow 0} \frac{\Pr(t \leq T < t + dt)}{dt} \quad (3)$$

Die gängigste Darstellungsform für die Verteilungen von Zeitdauern ist die im weiteren Verlauf der Modellentwicklung wichtige Hazard-Funktion  $h(t)$ . Alle drei im voraus beschriebenen Funktionen können auf die Hazard-Funktion zurückgeführt werden. Sie schätzt die ‚Wahrscheinlichkeit‘ bzw. das unmittelbare Risiko, dass der Eintritt eines Ereignisses in einer (kleinen) Zeitspanne zwischen  $t$  und  $t+dt$  zu erwarten ist – vorausgesetzt, dass das Ereignis bis zum Zeitpunkt nicht eingetreten ist. Damit werden nur die Wahrscheinlichkeiten für Individuen geschätzt, die zur aktuellen Risikogruppe (*risk set*) gehören, d.h. die bis Beginn des Intervalls  $t+dt$  überlebt haben. Im Gegensatz zur *Probability Density Function* – man beachte die Ähnlichkeit der Terme – spricht man jedoch bei der Hazard-Funktion von abhängiger (bedingter) Dichte, die sich aus der Restriktion zum *risk set* (s.o.) ergibt. In Anwendungsfällen, in denen wiederkehrende Ereignisse und abhängige Prozessdauern betrachtet werden, wird die Hazard-Funktion oft als Intensitätsfunktion bezeichnet.

$$h(t) = f(t) / S(t) = \lim_{dt \rightarrow 0} \frac{\Pr(t \leq T < t + dt | T \geq t)}{dt} \quad (4)$$

Die Interpretation von Hazard-Funktionen und Hazardraten ist aufgrund ihrer formalen Definition nicht unkompliziert und erfordert eine gewisse Vorstellungskraft des Betrachters. Hazardraten sind keine Wahrscheinlichkeiten im eigentlichen Sinne, sondern latente Variablen, die als Intensität des Übergangs von einem Zustand in einen weiteren aufgefasst werden können (Schneider, 1991). Je höher der Funktionswert, desto schneller findet im Durchschnitt der Übergang von Zustand A nach B



statt.

Abb. 5: Wesentliche Funktionen der *Survival Analysis*: Alternative, Verteilungsdarstellungen von T (Quelle: Verändert nach Hensher und Mannering (1994), S. 67)

Oft ist es von Interesse, nicht nur die Übergangswahrscheinlichkeit eines Zielzustandes (z.B.

eines künftigen Aktivitätenzweckes) zu betrachten, sondern eine Reihe verschiedener. Dazu können Modelle mit multiplen Übergangswahrscheinlichkeiten (*Konkurrierende Risiken*) geschätzt werden, die an anderer Stelle schon zur Modellierung der individuellen Zeitplanung eingesetzt wurden (Ettema *et al.*, 1995). Die Anwendung von Hazardmodellen in dieser Form kann als erster Schritt in eine neue Richtung der Simulation und Prognose von Tagesprogrammen angesehen werden, wird an dieser Stelle jedoch nicht weiter vertieft.

### 3. Deskriptive Analyse der Daten (Non-parametrisches Hazardmodell)

Bevor auf die eigentliche Entwicklung eines semi-parametrischen Modells eingegangen wird, die die Untersuchung der Determinanten für rhythmischen Verkehrsverhalten einschließt, erscheint zunächst eine Darstellung der univariaten Analyse der vorliegenden Intervall-Daten sinnvoll. Ziel ist hier, eine Schätzung der Survival Funktion (2) und der korrespondierenden Hazard-Funktion (4) zu erhalten – jeweils basierend auf den empirischen Beobachtungen zur rhythmischen Aktivitätenausübung.

Allgemein lässt sich für die (empirische) Survival Funktion folgendes formulieren:

$$\hat{S}(t) = \frac{\text{Anzahl der Beobachtungen } \geq t}{n} \quad t \geq 0 \quad (5)$$

Empirische Survivalraten berechnen sich üblicherweise über den sogenannten *Kaplan-Meier (KM)* oder *Produkt-Limit-Schätzer*, bei dem das Überleben von Prozessen bis zu beliebigen Punkten in der Zeit in eine Reihe von Schritten zerlegt wird. Die Schritte werden entweder durch die gemessenen Zeiten selber oder aber durch plausibel festgelegte Intervalle definiert.

Letzteres ist als Sterbetafelmethode (*life table method*) bekannt und bietet sich immer dann an, wenn – wie bei den *Mobidrive*-Intervalldaten – viele Beobachtungen vorliegen und Dauern exakt gemessen werden konnten. Eine so festgelegte Gruppierung der Ereigniszeiten in Intervalle (z.B. [0, 24 Stunden), [24, 48 Stunden), etc.) hat den Vorteil, dass die üblicherweise durch die KM-Methode generierten und wenig anschaulichen langen Tabellen mit verschiedensten statistischen Parametern auf ein gut darzustellendes Mass limitiert werden können. Eine Interpretation der Modellergebnisse wird so vereinfacht.

Fundamentales Prinzip des Sterbetafelansatzes ist die Annahme, das Ereignis- und Zensurzeitpunkte auf die jeweilige Intervallmitte zu datieren sind. Im *Mobidrive*-Fall berücksichtigen wir ausschliesslich unzensurierte Dauern, d.h. dass die Zeitintervalle vor der ersten Aktivität einer Person über die 42 Tage des Beobachtungszeitraums der Einfachheit halber ignoriert werden (siehe Abbildung 3, graue Schattierung). Die Wahl der Gruppierungsintervalle stellt einen mehr oder weniger willkürlichen Prozess dar – bei der empirischen Analyse der vorliegenden Dauerdaten schien eine Gruppierung der Daten in 24-Stunden-Intervalle (1 Tag) angemessen.

Die Survival Schätzung für das Intervall  $i$  via Sterbetafelmethode ist wie folgt definiert:

$$\hat{S}(t_i) = \prod_{j=1}^{i-1} (1 - q_j) \quad (6)$$

mit  $t_i$  = Startzeitpunkt des Intervalls  $i$   
 $q_i$  = Bedingte Wahrscheinlichkeit des Ereigniseintritts  
*(conditional probability of failure)*

Für  $i = 1$  und  $t_i = 0$  ergibt sich für den Survival-schätzer ein Wert von 1

Die *conditional probability of failure* ist dabei ein Schätzwert für die Wahrscheinlichkeit, dass ein Ereignis wie der Start einer Aktivität innerhalb des definierten Intervalls eintritt – wiederum vorausgesetzt, dass es bis  $t_i$  noch nicht beobachtet wurde. Diese Schätzung errechnet sich durch (Anzahl der Ereigniseintritte / Stichprobengrösse).

Die korrespondierende Hazard-Funktion ist gegeben durch

$$h(t_{im}) = \frac{d_i}{b_i \left( t_i - \frac{w_i}{2} - \frac{d_i}{2} \right)} \quad (7)$$

mit  $t_{im}$  = Intervallmitte  
 $b_i$  = Intervallbreite  
 $n_i$  = risk set am Beginn des Intervalls  
 $d_i$  = Anzahl der Ereignisse im Intervall  
 $w_i$  = Anzahl der zensierten Beobachtungen (hier: 0!)

Die Anwendungen des non-parametrischen Modells zur Analyse der Intervalle zwischen gleichartigen Aktivitäten bietet erste interessante Einblicke in

die Struktur der Zeitplanung und Aktivitätenerausübung über die sechs Wochen der *Mobidrive*-Erhebung (Abbildung 6). Die Intervallauern wurden für alle Aktivitätenkategorien (siehe Codierung im Anhang) untersucht. An dieser Stelle sollen nur ein Ausschnitt der (graphischen) Gesamtergebnisse präsentiert werden, bei dem besonderes Interesse den Freizeit- und Einkaufsaktivitäten gilt. Im Gegensatz zu den Aktivitäten Arbeit und Schule/Ausbildung bei denen von vornherein grosse Regelmässigkeit anzunehmen ist (vgl. Abbildung 6f), steht eine Untersuchung derer periodischen Struktur über längere Zeiträume noch aus.

Im allgemeinen ist für die Periodizität von Aktivitäten, die über die ganze Stichprobe in unregelmässigen zeitlichen Abständen ausgeübt werden, eine exponentiell-fallende Form der Survivalrate und eine konstante Hazardrate zu erwarten. Unsere Daten weisen dagegen erwartungsgemäss merklliche Strukturen rhythmischen Charakters auf.

Am deutlichsten ist dies bei den Intervallen zwischen Arbeitsaktivitäten von Vollzeitbeschäftigten (f), bei denen der Median der Ereigniszeiten (*Sterbezeitpunkt des Intervalls*) bei fast genau einem Tag liegt (Median=0.99861). Dies ist aufgrund der generellen zeitlichen Struktur von Arbeitszeiten nicht verwunderlich. Interessanter ist sicher die Tatsache, dass Freizeitaktivitäten – werden sie nach feineren Kategorien aufgeschlüsselt – einem festen zeitlichen Rhythmus folgen. Dies kann bei den Aktivitäten *Sport (aktiv)* und *Verein, Initiative etc.* (Abbildungen 6c und d) gezeigt werden, bei denen im Wochenrhythmus relative Maxima der Hazardraten, also der Intensitäten des Übergangs von einem Zustand in einen nächsten, zu beobachten sind.

Weniger sichtbar ist dies bei den weiteren Aktivitätszwecken, obwohl auch bei *Langfristigen Einkäufen* beim Ausgehen (*Lokal, Kino, Theater etc.*) ähnliche Tendenzen ohne Mühe festzustellen sind.

die Einflussgrößen der Periodizität des Verkehrsverhaltens getroffen worden, die über die Zeitabhängigkeit der Intervalldauer hinausgehen. Dies soll durch die Entwicklung eines semi-parametrischen proportionalen Hazardmodells oder Cox-Modells

scheinlichkeit von Ereignissen, sondern auch weitere Bestimmungsgrößen. Bei der Entwicklung eines angemessenen Modells für die Periodizität des Verkehrsverhaltens spielen – wie oben angedeutet – sozio-demographische Charakteristika, Öffnungszeiten von Einrichtungen, Arbeitszeitstrukturen oder persönliche Prämissen des Reisenden etc. eine Rolle. Im weiteren sollen die bis dato verfügbaren personen- und haushaltsbezogenen Determinanten des Mobilitätsverhaltens näher beleuchtet werden. Überlebenszeitmodelle, die den Einfluss solcher Determinanten schätzen können, fallen in die Klasse der *Proportional* oder *Parametric Hazard Models*.

In *Proportional Hazard Models* werden die zusätzlichen erklärenden Variablen als Funktion eines multi-dimensionalen Vektors definiert ( $X$ ), die multiplikatorisch auf die zugrundeliegende Hazard-Funktion (*baseline hazard*) wirkt (Abbildung 7). Bei dem in Abschnitt 2.) beschriebenen Basismodell wird dagegen angenommen, dass alle erklärenden (Ko-)Variablen den Wert 0 und damit keinen Einfluss auf die Intervalldauern besitzen.

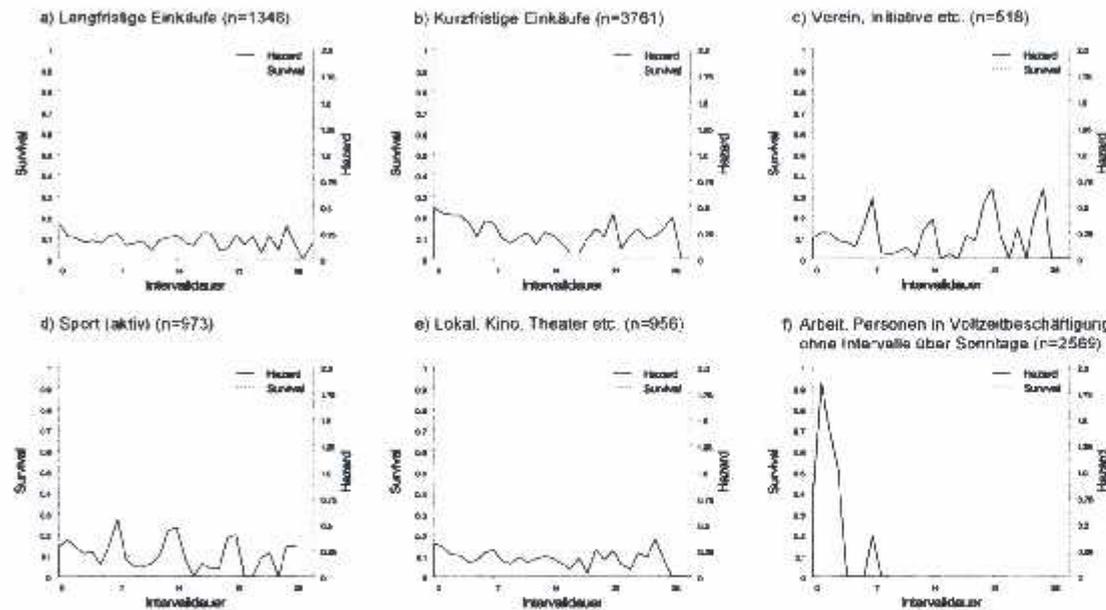


Abb. 6: Empirische Survival und Hazardraten ausgewählter Aktivitäten (Daten beider Fallstudienstädte)

#### 4. Erste Modellergebnisse mit dem semi-parametrischen Cox-Modell

Mit der in 3.) dargestellten Beschreibung der empirischen Daten sind noch keine Annahmen über

(Cox, 1972) geschehen.

##### Proportional Hazard Models

Hazardmodelle berücksichtigen im allgemeinen nicht nur die aktuelle Dauer eines Zustandes als wesentliche Determinante für die Eintrittswahr-

$$h(t|X) = h_0(t) g_0(X) = h_0(t) \exp(\beta X) \quad (8)$$

- mit  $X$  = Vektor der Kovariablen
- $\beta$  = Vektor der Schätzparameter
- $h_0(t)$  = baseline hazard

Beide Seiten logarithmiert ergibt

$$\log h_i(t) = \alpha(t) + \beta_i x_i + \dots + \beta_k x_k \quad (9)$$

Die Eigenschaften der Hazardfunktionen ändern sich somit proportional zur Ausprägung der Kovariablen, vorausgesetzt dass das Verhältnis verschiedener Hazardfunktionen eines Modells für be-

stimmte Sets von Kovariablen über die Zeit konstant bleibt ( $\approx h_1/h_2$ ). Anders ausgedrückt: Die Hazardrate eines Prozesses (eines Individuums) ist ein fester Anteil der Hazardrate aller anderen Prozesse.

$$\frac{h_i(t)}{h_j(t)} = \left\{ \exp \beta_1 (x_{i1} - x_{j1}) + \dots + \beta_k (x_{ik} - x_{jk}) \right\} \quad (10)$$

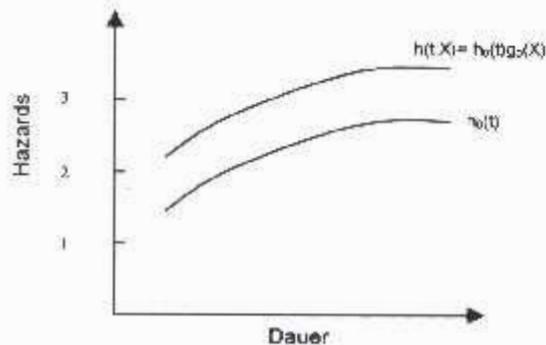


Abb. 7: Wirkung erklärender Variablen der Zeitnutzung auf das Hazard-Modell  
(Quelle: Eigene Darstellung nach Hensher und Mannering (1994), S. 68)

Eine Reihe von Modellen werden in der Klasse der *Proportional Hazards* zusammengefasst, bei denen der wesentliche Unterschied in den Verteilungsannahmen zur *baseline hazard* besteht (Annahmen zur *Probability Density Function* und  $\alpha(t)$ , Gleichungen 3 und 9). Die Gestalt von Hazard-Funktionen kann somit für bestimmte Prozesse zeitlicher Dauer verschiedene Formen annehmen (Abbildung 8). Möglich sind unter anderem eine monoton-steigende (a), eine U-förmige (b), eine mo-

noton-fallende (c) oder eine konstante Ausprägung (d). Bei monoton-ansteigenden Hazardraten (Weibullverteilung) kann man von Prozessen wie Entscheidungsfindungen ausgehen, deren Beendigung immer wahrscheinlicher wird, desto länger sie dauern. Prozesse, deren Beendigung mit fortschreitender Dauer unwahrscheinlicher wird, erzeugen eine monoton-fallende Hazardrate (ebenso Weibull). Eine konstante Ausprägung des Graphen (Exponentialverteilung) spiegelt dagegen Prozesse wieder, bei denen kein erkennbarer Zusammenhang zwischen dem Andauern und dem Abbruch des Vorgangs besteht. Letztlich sind auch Verläufe (wie bei b) vorstellbar, die keinen monotonen Charakter aufweisen (Log-logistische Verteilung).

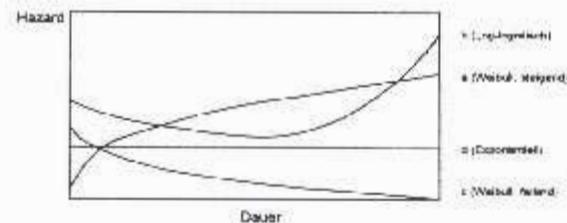


Abb. 8: Mögliche Ausprägungen von Hazard-Funktionen  
(Quelle: Nach Ettema et al. (1995), S. 102)

### Modellentwicklung

Als ein erster Schritt der Modellierung von rhythmischen Mustern des Verkehrsverhaltens soll ein Cox-Modell zur Anwendung kommen (vgl. Cox, 1972), das innerhalb der Analyse von Zeitdauern eine herausragende Stellung einnimmt. Schon an

anderer Stelle konnten sinnvolle Schätzungen zu Aktivitätendauern im Rahmen der Verkehrsverhaltensforschung angestellt werden (Hensher und Mannering, 1994; Mannering, Hammed und Kim, 1993; Sueyoshi, 1992). Bei weiteren Untersuchungen der Mobidrive-Daten steht eine Erweiterung des generellen Ansatzes um voll-parametrische Modelle an.

Der Grund für die Popularität des Cox-Regression-Ansatzes in vielen Forschungsfeldern besteht darin, dass zur *baseline hazard* keine expliziten Verteilungsannahmen getroffen werden müssen – deswegen der Begriff semi-parametrisch. Im Mittelpunkt des Interesses steht die Schätzung des Kovariantenparameters  $\beta$ . In der Literatur besteht Einigkeit, dass der Verzicht auf Verteilungsannahmen davor schützt, falsche Schlüsse aufgrund eines unpassenden Modells zur *baseline hazard* zu ziehen. Damit gilt das Cox-Modell generell als robuster verglichen mit voll-parametrischen Modellen. Ein weiterer Vorteil des Cox-Modells ist die Möglichkeit, zeitvariable Determinanten zu integrieren, also solche, die je nach Zeitpunkt der Beobachtung verschiedene Werte annehmen können (z.B. Wiedereinflüsse oder Erfahrungszuwächse aufgrund von kognitiven Lernprozessen). Desweiteren können sowohl diskrete als auch stetige Zeitdauern analysiert werden.

Erwähnt sei, dass die Entwicklung eines Cox-Modells allerdings auch Nachteile mit sich zieht, vor allem deswegen, weil sogenannte *heterogeneity effects* nicht berücksichtigt werden können. In Cox-Modellen wird implizit von der These ausgegangen, dass aufgrund der Proportionalität des Modells

Individuen oder Prozesse mit identischen Kovariablen auch identische Hazardraten aufweisen müssen. Werden keine Kovariablen berücksichtigt, existiert für die gesamte Stichprobe eine einzige Hazardrate. Dies ist offensichtlich eine unangemessene Voraussetzung, da sich Individuen und ihre Umwelt in vielerlei Hinsicht unterscheiden und die gemessenen / gewählten Kovariablen kaum das gesamte Spektrum der Heterogenität der Stichprobe abdecken können (vgl. Allison, 1995, 233ff.). Eine Nicht-Berücksichtigung der Heterogenität kann zu Verzerrungen der Schätzergebnisse von Modellen führen. Modelle, in denen Verzerrungen auftreten, weisen für die Hazardrate häufig fallende Werte auf, obwohl die wahre Hazardrate für keines der Individuen rückläufig ist. Allerdings wird an vielen Stellen – wie auch hier – ausdrücklich auf eine Integration verzichtet, da über den Charakter der Ungleichmässigkeit der Stichprobe oder der Umweltbedingungen Verteilungsannahmen getroffen werden müssten, die wiederum gewissen Unsicherheiten unterliegen. Zudem besteht auch im Cox-Ansatz genügend Raum für Zufallseffekte in den Zusammenhängen zwischen den Schätzwerten und den empirischen Dauerdaten.

Cox-Modelle basieren auf dem Prinzip der Maximierung von *partial likelihood*, das an dieser Stelle nicht vertieft erläutert werden soll. *Partial likelihood* ignoriert im wesentlichen den ersten Teil von (8) und behandelt den zweiten so, als ob es sich um eine herkömmliche Likelihood-Funktion handelt. Die Tatsache, dass im ersten Term (*baseline hazard*) zeitabhängige Informationen über den Schätzparameter  $\beta$  enthalten sein könnten, führt in der Regel

zu keinen grossen Verzerrungen des Modells, wird aber zumindest im weiteren überprüft.

Für die Schätzparameter gilt nach *partial likelihood*:

$$L(\beta) = \prod_{k=1}^n \left[ \frac{\exp(\beta X_k)}{\sum_{j \in R_k} \exp(\beta X_j)} \right] \quad (11)$$

mit  $L(\beta)$  = partial likelihood von  $\beta$   
 $\beta$  = Vektor der Schätzparameter  
 $n$  = Anzahl aller Dauern  
 $X$  = Vektor der Kovariablen

#### Exemplarische Modellschätzung

Die Vielzahl der verfügbaren Aktivitätenkategorien in *Mobidrive* bringt unbestreitbare Vorteile für den Detaillierungsgrad der Untersuchungen mit sich. Die Periodizität der Aktivitätenausübung divergiert zwischen den einzelnen Kategorien deutlich und lässt eine weitere Aggregation der Kategorien – beispielsweise in Pflichtaktivitäten versus Nicht-Pflichtaktivitäten – als zweitbeste Lösung erscheinen. Dies macht jedoch eine umfassende Darstellung aller Teilmodelle in einem Aufsatz problematisch. An dieser Stelle kann nur ein Ausschnitt der möglichen Ergebnisse gezeigt werden. Im folgenden soll exemplarisch ein Modell für die Intervalldauer der Aktivitäten *Kurzfristiger Einkauf* sowie *Sport (aktiv)* vorgestellt werden – zwei Kategorien, die bei der empirischen Untersuchung der Intervalle relativ deutliche Ergebnisse zur Periodizität gezeigt haben. Die Modelle wurden für die Befragten in den Städten Karlsruhe und Halle separat geschätzt.

Bevor auf die Schätzungen der entwickelten Regressionsmodelle eingegangen wird, sollen einige Anmerkungen zu den Spezifika der Statistik gemacht werden:

Zunächst ist darauf hinzuweisen, dass die Überprüfung der grundlegenden Annahme zur Proportionalität der *baseline hazard* am Beginn jeder Schätzung mit Cox-Modellen stehen sollte. Dazu existieren eine Reihe von unterschiedlichen Methoden (vgl. Kleinbaum, 1996, 129ff.), die das grundlegende theoretische Konstrukt des Cox-Modells, d.h. die multiplikative Wirkung eines zeit-unabhängigen Kovariablenvektors auf die *baseline*, bestätigen bzw. widerlegen sollen. Zur Evaluierung der hier vorgestellten Modelle wurde eine übliche graphische Methode gewählt, nämlich der Vergleich von geschätzten *log-log survival curves* über die verschiedenen Ausprägungen der gewählten Kovariablen. *Log-log survival curves* sind einfache Transformationen der geschätzten Survivalfunktionen, wobei deren Ergebnisse logarithmiert und anschliessend mit dem Wert zwei multipliziert werden (mathematisch:  $-\ln(-\ln \hat{S})$ ). Soll die Annahme der proportionalen Hazards bestätigt werden, müssen die Funktionsgraphen der *log-log survival curves* möglichst parallel verlaufen. Auf eine eingehende Herleitung der Teststatistik wird an dieser Stelle verzichtet. Nach der separaten Überprüfung der gewählten Kovariablen gehen wir davon aus, dass die gezeigten Modelle die Proportionalitätsannahme erfüllen.

Im jeweils oberen Teil der folgenden Tabellen sind die Koeffizientenschätzungen und die sogenannte *risk ratio* aufgeführt. Die Besonderheit von *partial likelihood* Modellen besteht darin, dass keine Schätzungen des konstanten Glieds der Gleichung

(*intercept*) durchgeführt werden. Die Konstante ist Teil der logarithmierten *baseline* Funktion (vgl. Gleichung 9).

Von grösstem Interesse ist die *risk* oder *hazard ratio*, die für Cox-Modelle durch  $e^{\beta}$  errechnet werden kann. Sie ist ein Indikator dafür, wie sich die Hazardrate eines Individuums (eines Prozesses) gegenüber anderen gestaltet, d.h. sie stellt den quantitativen Wert für die multiplikatorische Wirkung des Modells dar (siehe oben, Abbildung 8). Für Indikator- bzw. Dummyvariablen (beispielsweise Geschlecht: weiblich=0, männlich=1) kann die *risk ratio* als das Verhältnis der geschätzten Hazardrate für Individuen mit dem Wert 1 an der Hazardrate für Individuen mit dem Wert 0 interpretiert werden. Beispielsweise ist die geschätzte *risk ratio* für die Befragten mit mehr als 30 Stunden wöchentlicher Arbeitszeit beim Modell für kurzfristige Einkäufe (Karlsruhe) gleich 0.62. Dies bedeutet, dass der Hazard für den Ereigniseintritt (Start der folgenden gleichartigen Aktivität) nur bei etwa 66% derjenigen mit weniger als 30 Wochenarbeitsstunden liegt und damit die Intervalle zwischen den Einkaufsaktivitäten um etwa ein Drittel länger sind – unter statistischer Berücksichtigung der anderen Kovariablen. Für kontinuierliche Variablen ist eine sinnvolle Statistik durch die Subtraktion von 1 und einer anschließenden Multiplikation mit 100 zu erreichen. Dies ergibt die geschätzte prozentuale Änderung der Hazardrate für jeden Anstieg der Kovariable um eine Einheit. Für die Variable *Anzahl PKW im Haushalt* im gleichen Modell ergibt sich eine *risk ratio* von 0.85, was zu einer um etwa 15 Prozentpunkte 100(0.85-1) niedrigeren Hazardrate für jedes zusätzliche Automobil im Haushalt führt.

	Karlsruhe		Halle	
	Parameterschätzung ( $\beta$ )	Risk Ratio	Parameterschätzung ( $\beta$ )	Risk Ratio
<b>Personenbezogen</b>				
Ist männlich	-0.033	1.03	-0.139	0.67 *
Alter	0.026	1.03 *	0.051	1.05 *
Alter (quadrat)	-0.006	1.00 *	-0.009	1.00 *
Hat PKW-Führerschein	-0.147	0.85 *	0.103	1.11
Ist Vereinsmitglied	-0.009	1.00	-0.093	0.91
Ist verheiratet	0.033	1.03	0.047	1.05
Ist Vater / Mutter	0.348	1.27 *	0.195	1.11
<b>Haushaltsbezogen</b>				
Anzahl HH-Mitglieder	-0.086	0.92 *	-0.035	0.96
Anzahl PKW	-0.162	0.85 *	-0.060	0.94
Netto-Einkommen > 2000€M / HH-Mitglied	0.385	1.36 *	0.211	1.24 *
<b>Beschäftigungstatus</b>				
Arbeitszeit > 30 h	-0.475	0.62 *	-0.257	0.77 *
N		2024		1737
LogLikelihood(0)		26781.452		23454.022
LogLikelihood(1)		20505.906		22361.440
ChiSquare (DF=15)		215.548		72.535
R <sup>2</sup>		0.101		0.041

\* Statistisch signifikant bei 95%-Konfidenzintervall

Tab. 1: Parameterschätzung für semi-parametrischen Cox-Ansatz (Aktivität: Kurzfristiger Einkauf – Positiver Proportionalitätstest)

	Karlsruhe		Halle	
	Parameterschätzung ( $\beta$ )	Risk Ratio	Parameterschätzung ( $\beta$ )	Risk Ratio
<b>Personenbezogen</b>				
Ist männlich	0.332	1.45 *	0.319	1.38 *
Alter	0.013	1.01	0.041	1.04
Alter (quadrat)	-0.000	1.00	-0.000	1.00
Hat PKW-Führerschein	0.498	1.55 *	0.638	1.71 *
Ist Vereinsmitglied	0.622	1.86 *	1.089	2.97 *
Ist verheiratet	0.012	1.01	-0.410	0.66
Ist Vater / Mutter	-0.280	0.77	-0.180	0.83
<b>Haushaltsbezogen</b>				
Anzahl HH-Mitglieder	0.051	1.06	0.243	1.28 *
Anzahl PKW	0.025	1.02	0.110	1.13
Netto-Einkommen > 2000€M / HH-Mitglied	-0.031	0.97	0.429	1.54 *
<b>Beschäftigungstatus</b>				
Arbeitszeit > 30 h	-0.278	0.76 *	-0.742	0.48 *
N		872		901
LogLikelihood(0)		7417.353		2842.796
LogLikelihood(1)		7321.580		2753.406
ChiSquare (DF=11)		95.613		47.26
R <sup>2</sup>		0.133		0.252

\* Statistisch signifikant bei 95%-Konfidenzintervall

Tab. 2: Parameterschätzung für semi-parametrischen Cox-Ansatz (Aktivität: Sport (aktiv) – Positiver Proportionalitätstest)

Bei der inhaltlichen Interpretation der Schätzergebnisse soll nur auf einige interessante Punkte eingegangen werden: Für die Modellschätzung beider Aktivitätskategorien ist anzunehmen, dass die Auswahl der Kovariablen nur bedingten Einfluss auf die Intervalllängen hat (niedriges  $R^2$ ). Die Streuung im Modell für die Aktivität *Sport (aktiv)* kann für beide Städte jedoch erheblich stärker auf die gewählten Determinanten zurückgeführt werden. Insgesamt ist die Auswahl der Kovariablen aufgrund von Ad-hoc-Annahmen zu den Einflussgrößen von Mobilität und der Proportionalitätstests getroffen worden – sie wird in weiteren Modellschätzungen durch Tests zur Signifikanz einzelner Variablen und ihrer Wechselwirkungen sowie einer Verfeinerung der konzeptionellen Basis modifiziert.

Generell lässt sich feststellen, dass zwischen den beiden Fallstudienstädten wesentliche Unterschiede bei den Einflussgrößen auf das zeitliche Verhalten der Befragten besteht. Eine grundlegende – allerdings eher quantitative als qualitative – Erklärung dafür sind die etwa 20% mehr berichteten Wegen pro Tag in der Karlsruher gegenüber der Hallenser Stichprobe über den Sechs-Wochen-Zeitraum. Anders ausgedrückt: Entwickelt man ein bivariates Cox-Regressionsmodell für die Aktivität *Kurzfristiger Einkauf*, in dem ausschliesslich eine Indikatorvariable für die Fallstudienstadt berücksichtigt wird, liegt die Hazardrate für Karlsruhe um etwa 15 Prozentpunkte über der von Halle.

Die berücksichtigten Einflussgrößen unterscheiden sich insbesondere beim Modell für die Aktivität *Kurzfristiger Einkauf*. Während in der Karlsruher Stichprobe kürzere Intervalle für vermeintlich (auto-)mobile Verkehrsteilnehmer festzustellen

sind (Führerscheinbesitz / Autobesitz des Haushalts), ist dies in Halle nicht der Fall. Dort sind eher die längeren Intervalle der Männer auffällig. Parallele Tendenzen können in beiden Städten bei der Signifikanz der Vollzeitberufstätigkeit (deutlich längere Intervalle zwischen den Einkäufen) und beim Einfluss des höheren Haushaltseinkommens (höhere Frequenz der Einkaufsaktivitäten) beobachtet werden.

Beim Modell für die Aktivität Sport sind die Variablen Geschlecht, Führerscheinbesitz und Vereinsmitgliedschaft signifikant für beide Städte. Insbesondere Männer und – wie zu erwarten war – Befragte, die in Vereinen organisiert sind, weisen kürzere Intervalle zwischen den Sportaktivitäten auf. Interessant stellt sich auch die ‚grosse sportliche Aktivität‘ der Führerscheininhaber dar, die in beiden Städten gegenüber den anderen Verkehrsteilnehmerinnen und –teilnehmern eine fast zweimal höhere Intensität dieser Aktivitätsausübung aufweisen. Deutlich fallen auch die *risk ratios* der Vollzeitbeschäftigten (mehr als 30 Stunden Wochenarbeitszeit) aus, die wesentlich unter denen der Vergleichsgruppe liegen.

## 5. Methodischer und inhaltlicher Ausblick

Die hier entwickelte Methodik zur Analyse der Periodizität des Verkehrsverhaltens stellt eine sinnvolle Erweiterung des Methodenpools innerhalb der aktivitäten-basierten Verkehrsforschung dar. Das

Basiskonzept bietet eine Reihe von Ansatzpunkten zur Modifizierung, worauf abschliessend in Kürze eingegangen werden soll:

*Entwicklung eines voll-parametrischen Modells; Integration von Verteilungsannahmen zur baseline hazard:*

Wie oben beschrieben, bietet der Cox-Ansatz eine robuste Modellschätzung, ohne dass Verteilungsannahmen zum zeitlichen Verlauf der *baseline hazard* angestellt werden müssen. Eine weitere theoretische und konzeptionelle Spezifizierung des Charakters von rhythmischen Strukturen des Verkehrsverhaltens, die von den dargestellten Minimalannahmen abrückt, lässt die Anwendung eines voll-parametrischen Modells sinnvoll erscheinen. Dabei versprechen Weibull-Verteilungen, die zu monotonsteigenden Hazardfunktionen führen (vgl. Abbildung 4 und 8), grösste Realitätsnähe.

*Unbeobachtete bzw. bisher nicht unberücksichtigte Heterogenität innerhalb der Stichprobe:*

In einer Weiterentwicklung des Ansatzes sollen *heterogeneity effects* berücksichtigt werden. Dabei ist nicht nur die allgemeine unbeobachtete Heterogenität zwischen den Befragten von Interesse, sondern auch nicht berücksichtigte Effekte vorausgegangener Dauern von Intervallen (sogenanntes *state dependence*, vgl. Hensher und Mannering, 1994, 72ff.).

*Competing Risks* und Anwendung der Modelle auf Rhythmen zwischen komplexeren Verhaltensmustern:

Nicht nur die Verbesserung der zugrundeliegenden Modellannahmen steht im Mittelpunkt der weiteren Arbeit, sondern auch die konzeptionelle Erweiterung der Modelle um Aspekte wie konkurrierende Ereignisausgänge (siehe oben) oder die Berücksichtigung mehrdimensionaler Verhaltensmuster als Analysegegenstand.

Letztlich bietet der vorgestellte Ansatz Anknüpfungspunkte mit parallelen Entwicklungen in der aktivitäten-basierten Verkehrsforschung. Insbesondere im Bereich der umfassenden Mikrosimulation des Aktivitätenwahl- und Zeitplanungsverhaltens von Personen und Haushalten (*scheduling*) kann eine Integration der Modellschätzungen zur Verbesserung der Werkzeuge führen.

## Abschließende inhaltliche Bemerkungen

Inhaltlich deutet die Auswahl der empirischen Ergebnisse darauf hin, dass Zeitplanung und Verkehrsverhalten zeitlichen Mustern folgen, die in weitem Teilen täglichen, bei einer Reihe von Einkaufs- bzw. Freizeitaktivitäten jedoch wöchentlichen Strukturen entsprechen. Eingehendere Analysen werden zeigen, inwieweit routinisiertes Verhalten der Verkehrsteilnehmer für diese Phänomene verantwortlich zeichnen und welchen Anteil Umwelt-rhythmen an der Periodizität des Verhaltens haben. Im wesentlichen bestätigen die Ergebnisse des semi-parametrischen Cox-Modells die generellen Annahmen zu den Einflussgrössen von Raum-Zeit-Verhalten:

*Sozio-demographische Attribute der Reisenden:*

Die Sozio-Demographie der Reisenden und insbesondere deren Rolle im Haushaltszusammenhang (Haushaltszusammensetzung, Kinder, Fahrzeugverfügbarkeit, Aufgabenverteilung im Haushalt etc.) bestimmen nicht nur die quantitative Nachfrage nach Mobilitätsdienstleistungen sondern auch die Struktur der Zeitplanung ausserhäuslicher Aktivitäten. Die vorgestellte Analyse erschliesst zudem interessante Resultate über die unterschiedlichen Bestimmungsgrössen des zeiträumlichen Alltagsverhaltens in Ost und West.

*Konsum und Freizeitstile:*

Der Charakter der individuellen Konsum- und Freizeitgestaltung von Personen und Haushalten wie Einkaufsgewohnheiten, Lebensstile, Einteilung des disponiblen Einkommens, Selbstverpflichtungen, Gruppen- und Vereinszugehörigkeit oder Einstellungen gegenüber Verkehrsmitteln sind bedeutende Faktoren der Mobilität. Sie treten insbesondere deswegen in den Vordergrund, weil der Freizeit- und Einkaufsverkehr einen grösseren Anteil am Gesamtverkehrsaufkommen erlangt. Die Verknüpfung der hier vorgestellten Daten mit den zusätzlich erhobenen Informationen zu den Werthaltungen und Einstellungen der Befragten verspricht aufschlussreiche Analysen zu alternativen Determinanten der Periodizität des Verhaltens.

Die Autoren danken den an Mobidrive beteiligten Projektpartnern A. Zimmermann und T. Haupt (PTV AG), G. Rindsfuser, T. Wehmeier und K.J. Beckmann (Institut für Stadtbauwesen, RWTH Aachen) sowie insbesondere den IVT-Kollegen A. König und R. Schlich für ihre Unterstützung bei der Vorbereitung dieses Aufsatzes.

**Anhang****Aktivitätenkategorisierung**

*Standardkategorisierung (vgl. KONTIV o.ä.)*

- 1 Jemanden bringen / abholen
- 2 Dienstleistungen / Erledigungen
- 3 Arbeitsbezogen
- 4 Schule / Ausbildung
- 5 Arbeit
- 6 Kurzfristiger Einkauf
- 7 Langfristiger Einkauf
- 8 Freizeit
- 9 Sonstiges
- 10 Nach Hause

*City:mobil-Kategorisierung (vgl. Götz, Jahn und Schuitz, (1997))*

- 1 Arbeit
- 2 Beruflicher Zusammenhang
- 3 Ausbildungsplatz
- 4 Fort-/Weiterbildung
- 5 Versorgung von Kindern oder anderen Familienangehörigen
- 6 Einkauf des täglichen Bedarfs
- 7 Einkauf grosser Vorräte und besonderer Anschaffungen
- 8 Einkaufsbummel
- 9 Erledigung von Pflichten (Bank, Post, ...)
- 10 Besuch bei Verwandten
- 11 Treffen mit Freunden
- 12 Verein, Bürgerinitiative, Kirche etc.
- 13 Arzt, Zahnarzt, Friseur
- 14 Autopflege, Tanken etc.
- 15 Ausübung von Sport
- 16 Grösserer Ausflug in die Natur
- 17 Spaziergang, Spazierfahrt
- 18 Kurzurlaub, Wochenendausflug
- 19 Garten, Gartenhaus
- 20 Etwas neues sehen (Stadt, Sightseeing etc.)
- 21 Disko, Lokal, Kneipe, Restaurant
- 22 Nach Hause
- 23 Sonstiges

## Literatur

- [1] Allison, P.D. (1995)  
*Survival Analysis using the SAS system – A practical guide*, SAS Institut Inc., Cary.
- [2] Axhausen, K.W., A. Zimmermann, S. Schönfelder, G. Rindsfuser und T. Haupt (2000)  
Observing the rhythms of daily life: A six-week travel diary, *Arbeitsberichte Verkehrs- und Raumplanung*, **25**, Institut für Verkehrsplanung und Transporttechnik, ETH, Zürich.
- [3] Bhat, C.R. (1996a)  
A hazard based duration model of shopping activity with nonparametric baseline specification and nonparametric control for unobserved heterogeneity, *Transportation Research B*, **30** (3) 189-207.
- [4] Bhat, C.R. (1996b)  
A generalized multiple durations proportional hazard model with an application to activity behaviour during the evening work-to-home commute, *Transportation Research B*, **30** (6) 465-480.
- [5] Cox, D.R. (1972)  
Regression models and life tables, *Journal of the Royal Statistical Society*, **28B**, 186-220.
- [6] Cox, D.R. (1984)  
*Analysis of Survival Data*, Chapman and Hall, New York.
- [7] Ettema, D., A. Borgers und H. Timmermans (1995)  
Competing risk hazard model of activity choice, timing, sequencing, and duration, *Transportation Research Record*, **1493**, 101-109.
- [8] Götz, K., T. Jahn, und I. Schultz. (1997)  
Mobilitätsstile: Ein sozial-ökologischer Untersuchungsansatz, *Forschungsbericht Stadtverträgliche Mobilität*, **7**, Forschungsverbund City:mobil, Frankfurt am Main.
- [9] Hägerstrand, T. (1970)  
What about people in regional science?, *Papers of the Regional Science Association*, **24** (1) 7-21.
- [10] Hamed, M.K. und F. Mannering (1992)  
Modelling travellers' post-work activity involvement: toward a new methodology, *Transportation Science*, zur Veröffentlichung eingereicht.
- [11] Hamed, M.K., S.G. Kim und F. Mannering (1992)  
A note on travelers' home-stay duration and the efficiency of proportional hazard models, Working Paper, University of Washington, Seattle.
- [12] Hanson, S. und Hanson, P. (1981a)  
The travel-activity patterns of urban residents: Dimensions and relationships to sociodemographic characteristics, *Economic Geography*, **57**, 332-347.
- [13] Hanson, S. und Hanson, P. (1981b)  
The impact of married women's employment on household travel patterns: A Swedish example, *Transportation*, **10**, 165-183.
- [14] Hanson, S. und K.O. Bumett (1981)  
Understanding complex travel behavior: Measurement issues, in P.R. Stopher, A.H. Meyburg und W. Brög, *New Horizons in Travel-Behaviour Research*, 207-230, Lexington, Massachusetts/Toronto.
- [15] Hensher, D.A. und Mannering, F.L. (1994)  
Hazard-based duration models and their application to transport analysis, *Transportation Reviews*, **14** (1) 63-82.
- [16] Hosmer, D. W. und S. Lemeshow (1999)  
*Applied Survival Analysis: Regression Modeling of Time to Event Data*, Wiley, New York.
- [17] Jones, P.M. (1981)  
Activity approaches to understanding travel behaviour, in P.R. Stopher, A.H. Meyburg und W. Brög, *New Horizons in Travel-Behaviour Research*, 253-266, Lexington, Massachusetts/Toronto.
- [18] Kalbfleisch, J.D. und R.L. Prentice (1980)  
*The Statistical Analysis of Failure Time Data*, Wiley, New York.
- [19] Kitamura, R. (1988)  
An evaluation of activity-based travel analysis, *Transportation*, **15** (1) 9-34.
- [20] Kleinbaum, D. G. (1996)  
*Survival analysis : A Self-learning Text*, Springer, New York.
- [21] Mahmassani, H.S. (1988)  
Some comments on activity-based approaches to analysis and prediction of travel behavior, *Transportation*, **15** (1) 35-40.

- [22] Mannering, F. und H. Hamed (1990)  
Occurrence, frequency and duration of commuters' work-to-home departure delay, *Transportation Research B*, **24** (2) 99-109.
- [23] Mannering, F., E. Murakami und S.G. Kim (1992)  
Models of travelers' activity choice and home-stay duration: analysis of functional form and temporal stability, Working Paper, University of Washington, Seattle.
- [24] Niemeier, D.A. und J. Morita (1994)  
Duration of trip-making activities by men and women: a survival analysis, Presented at the 73<sup>rd</sup> Annual Meeting of the Transportation Research Board
- [25] Oh, L. (2000)  
Identification of the causal factors of activity duration dynamics using hazard based duration models, Working paper, Centre for Transport Studies, Imperial College of Science, Technology and Medicine, London.
- [26] Pas, E.I. (1983)  
A flexible and integrated methodology for analytical classification of daily travel-activity behavior, *Transportation Science*, **17** (4) 405-429.
- [27] Pas, E.I. und A.S. Harvey (1997)  
Time use research and travel demand analysis modelling, in P.R. Stopher und M.E.H Lee-Gosselin (Hrsg.) *Understanding Travel Behaviour in an Era of Change*, 315-338, Pergamon, Oxford.
- [28] Reader, S. und F.R. McNeill (1999)  
Hazard-rate modelling of store switching behaviour, *Environment and Planning A*, **31** (8) 1353-1370.
- [29] Recker, W.W., M.G. McNally und G.S. Root (1983)  
Application of pattern recognition theory to activity pattern analysis, in S. Carpenter und P.M. Jones (Hrsg.) *Recent Advances in Travel Demand Analysis*, 434-449, Gower, Aldershot.
- [30] Schlittgen, R. und B.J.H. Streitberg (1999)  
*Zeitreihenanalyse*, Oldenburg, München.
- [31] Schneider, H. (1991)  
*Verweildaueranalyse mit GAUSS*, Campus, Frankfurt/New York.
- [32] Shapcott, M. und P. Steadman (1978)  
Rhythms of urban activity, in T. Carlslein, D. Park und N. Thrift (Hrsg.) (1978) *Human Activity and Time Geography*, 49-74, Arnold, London.
- [33] Sueyoshi, G.T. (1992)  
Semiparametric proportional hazards estimation of competing risks models with time varying covariates, *Journal of Econometrics*, **51**, 25-58.

#### Dipl.-Ing. Stefan Schönfelder

Institut für Verkehrsplanung und Transporttechnik, Straßen- und Eisenbahnbau (IVT), Eidgenössische Technische Hochschule (ETH) Zürich  
ETH-Hönggerberg, CH-8093 Zürich  
schoenfelder@ivt.baug.ethz.ch

#### Prof. Dr.-Ing. Kay W. Axhausen

Institut für Verkehrsplanung und Transporttechnik, Straßen- und Eisenbahnbau (IVT), Eidgenössische Technische Hochschule (ETH) Zürich  
ETH-Hönggerberg, CH-8093 Zürich  
axhausen@ivt.baug.ethz.ch