



**FORSCHUNG**  
MOBILITÄT  
TRANSPORT  
VERKEHR

# **Laufende Hochrechnung der Kfz- Fahrleistungen auf der Basis von Hauptuntersuchungsdaten**

im Auftrag des Bundesministeriums für Verkehr, Bau  
und Stadtentwicklung  
FE-Nr.: 70.0832/2008

## **Schlussbericht Band 1**

### ***Hochrechnungsverfahren***

Prof. Dr. Heinz Hautzinger  
Dipl.-Sozialwiss. Mario Fuchs  
Dr. Jochen Schmidt  
Dr. Wilfried Stock

Institut für angewandte Verkehrs-  
und Tourismusforschung e.V.

[www.ivt-verkehrsforschung.de](http://www.ivt-verkehrsforschung.de)

Heilbronn / Mannheim, den 31.05.2010

## Inhaltsverzeichnis

<b>1</b>	<b>Ausgangslage und Aufgabenstellung .....</b>	<b>5</b>
<b>2</b>	<b>Methodisches Rahmenkonzept der Untersuchung .....</b>	<b>9</b>
2.1	Fahrleistung als Gegenstand der empirischen Forschung .....	9
2.1.1	Grundlegende Definitionen und Konzepte .....	9
2.1.2	Spezielle Fahrleistungskennzahlen .....	10
2.2	Methoden zur Gewinnung von statistischen Informationen über Fahrleistungen.....	11
2.2.1	Primärstatistische Ansätze .....	12
2.2.2	Sekundärstatistische Ansätze .....	13
2.2.3	Eignung der verschiedenen Verfahren zur Fahrleistungsermittlung .....	14
<b>3</b>	<b>Statistisches Grundmodell der Fahrleistungen von Kraftfahrzeugen.....</b>	<b>16</b>
3.1	Periodenbezogene Fahrleistung von Einzelfahrzeugen .....	16
3.2	Periodenbezogene Fahrleistung eines Fahrzeugkollektivs .....	18
3.2.1	Anmeldequote und mittlerer Bestand angemeldeter Fahrzeuge.....	18
3.2.2	Periodenbezogene Fahrleistungskennzahlen .....	19
3.3	Bisherige Fahrleistung: Tachostand von Fahrzeugen .....	21
3.3.1	Bisherige Lebensdauer und bisherige Fahrleistung von Fahrzeugen.....	21
3.3.2	Bisherige versus aktuelle Fahrleistung von Fahrzeugen .....	22
3.3.3	Stichtagsbezogener mittlerer Tachostand von Fahrzeugen.....	24
3.3.4	Periodenbezogener mittlerer Tachostand von Fahrzeugen .....	25
3.4	Das Dilemma: Schätzung periodenbezogener Fahrleistungen auf der Basis stichtagsbezogener Tachostandsdaten .....	25
3.4.1	Periodenbezogene Kennzahlen als Zielgrößen der Verkehrsstatistik .....	25
3.4.2	Tachostandsdaten als empirisches Material der Verkehrsstatistik .....	27
3.4.3	Tachostand als Hilfsvariable für eine Fahrleistungsschätzung .....	28
<b>4</b>	<b>Prozessdaten aus Kfz-Hauptuntersuchungen als Grundlage einer Fahrleistungsstatistik .....</b>	<b>30</b>
4.1	Prozessdaten als empirische Basis statistischer Berichtssysteme .....	30

4.2	Technische Überwachungsinstitutionen als Lieferanten von Prozessdaten für Zwecke einer Fahrleistungsstatistik.....	31
4.3	Fahrzeug- und prüfterminbezogene Merkmale des HU-Rohdatensatzes .....	32
<b>5</b>	<b>Methodische Grundsatzfragen der Eignung von HU-Daten als Basis einer Fahrleistungsstatistik .....</b>	<b>34</b>
5.1	Die methodischen Probleme im Überblick.....	34
5.2	Repräsentativität der bei den datenliefernden Prüforganisationen untersuchten Fahrzeuge für alle untersuchten Fahrzeuge.....	35
5.2.1	Strukturvergleich zwischen HU-Stichprobe und FU-Statistik .....	35
5.2.2	Validität der HU-Daten bezüglich der Untersuchungsintervalle .....	36
5.3	Repräsentativität der in einer Periode geprüften Fahrzeuge für den betreffenden Fahrzeugbestand insgesamt .....	38
5.3.1	Beschränkung auf geprüfte Fahrzeuge als Coverage-Problem.....	38
5.3.2	Strategien zur Reduzierung von Verzerrungen .....	39
<b>6</b>	<b>Alternative Modellansätze für die Schätzung periodenbezogener Fahrleistungen aus Tachostandsdaten .....</b>	<b>40</b>
6.1	Kohortenübergreifende Schätzung auf der Basis von Querschnittsdaten.....	40
6.2	Regressionsanalytische Schätzung auf der Basis von Querschnittsdaten .....	42
6.3	Kohortenspezifische Schätzung auf der Basis von Längsschnittdaten .....	44
6.4	Endogene Schätzung auf der Basis von Querschnittsdaten: Die Durchschnittswertmethode .....	45
6.4.1	Grundgedanke.....	45
6.4.2	Erste kritische Diskussion.....	46
6.5	Resümee.....	49
<b>7</b>	<b>Empirische Prüfung des Konzepts der endogenen Schätzung .....</b>	<b>50</b>
7.1	Überprüfung anhand der Fahrleistungserhebung 2002.....	50
7.2	Überprüfung anhand der Erhebung von Fahrleistung und Kraftstoffverbrauch im Rahmen des Deutschen Mobilitätspanels .....	53
7.3	Zusammenfassende Bewertung der Ergebnisse der empirischen Überprüfung der Durchschnittswertmethode .....	55

<b>8</b>	<b>Verfahren zur Aufbereitung der HU-Rohdaten für Zwecke einer Fahrleistungsstatistik</b>	<b>56</b>
8.1	Vereinheitlichung von Formaten .....	56
8.2	Bereinigung um Dubletten und Fälle mit ungültigen Variablenwerten.....	56
8.3	Zusammenführung der einzelnen Rohdatendateien und Bereinigung um Fälle mit ungültigen Angaben zum Tachostand .....	57
8.4	Ergänzung der HU-Rohdaten um fahrzeugtechnische Merkmale .....	58
8.4.1	Merkmal „Fahrzeugart“ .....	58
8.4.2	Merkmal „Antriebsart“ .....	60
8.4.3	Leistungs- und Größenmerkmale .....	61
8.4.4	Grenzen der Datenergänzung .....	61
8.5	Ergänzung der HU-Rohdaten um das Merkmal Fahrzeugalter .....	62
8.5.1	Fahrzeugalter zum Zeitpunkt der Hauptuntersuchung.....	62
8.5.2	Fahrzeugalter im Bezugsjahr der Hochrechnung.....	63
8.6	Plausibilitätsprüfungen bezogen auf das Merkmal Tachostand .....	64
8.6.1	Abschneidegrenzen für Tachostände.....	64
8.6.2	Korrektur von Verzerrungen durch Tachoüberläufe .....	67
8.6.3	Zusammenfassung des Bereinigungs- bzw. Korrekturverfahrens .....	70
8.7	Auswirkungen der Datenbereinigung auf das Datenvolumen .....	70
<b>9</b>	<b>Verfahren zur Hochrechnung kalenderjahrbezogener Fahrleistungen auf der Basis von HU-Daten.....</b>	<b>72</b>
9.1	Anforderungen an das Hochrechnungsverfahren.....	72
9.2	Prinzip der Übereinstimmung von HU-Prüfjahr und Bezugsjahr der Hochrechnung	74
9.3	Grundstruktur des Hochrechnungsverfahrens.....	74
9.3.1	Hochrechnungsrelevante Variablen .....	74
9.3.2	Schätzfunktion für den kalenderjahrbezogenen Fahrleistungstotalwert .....	76
9.3.3	Fahrleistungsschätzung für Teilgruppen von Fahrzeugen .....	76
9.4	Hochrechnungsgruppen: Nachträgliche Schichtung der HU-Stichprobe .....	77
9.4.1	Schichtungsmerkmal „Fahrzeugart“ .....	77
9.4.2	Schichtungsmerkmal „Zulassungsjahr“ bzw. „Fahrzeugalter“ .....	78
9.4.3	Weitere Merkmale für die Bildung von Hochrechnungsgruppen.....	81
9.4.4	Der Schichtungsplan im Überblick .....	81

9.5	Korrektur des Coverage-Fehlers: Konzept der „neutralen“ Fahrleistung .....	82
9.5.1	Ursachen und Folgen des Coverage-Fehlers .....	82
9.5.2	Folgerungen für das Hochrechnungsverfahren.....	84
9.5.3	Grundgedanke des Korrekturverfahrens .....	85
9.5.4	Verzerrungskorrektur nach dem Konzept der „neutralen“ Fahrleistung .....	86
9.5.5	Bestimmung der Anteilswerte geprüfter und ungeprüfter Fahrzeuge .....	90
9.6	Hochrechnungsfaktoren: Gesamtzahl der Fahrzeugmonate und mittlerer Fahrzeugbestand im Bezugsjahr der Hochrechnung .....	91
9.6.1	Bestimmung der Gesamtzahl der Fahrzeugmonate im Bezugsjahr.....	91
9.6.2	Schätzung des mittleren Fahrzeugbestands im Bezugsjahr .....	93
9.7	Grenzen des Hochrechnungsverfahrens .....	94
<b>10</b>	<b>Hochrechnungsergebnisse.....</b>	<b>96</b>
10.1	Hochgerechnete Jahresfahrleistungen 2006 und 2007.....	96
10.2	Vergleich mit Ergebnissen der DIW-Modellrechnungen.....	96
10.2.1	Grad der Übereinstimmung .....	96
10.2.2	Mögliche Ursachen für Ergebnisunterschiede.....	98
10.3	Beispiele zur Ergebnisdarstellung .....	99
10.3.1	Jahresfahrleistung von Teilgruppen unterschiedlicher Fahrzeugarten .....	99
10.3.2	Mittlere Jahresfahrleistungen 2006 und 2007 nach Fahrzeugart und -alter100	
<b>11</b>	<b>Kritische Diskussion des entwickelten Hochrechnungsverfahrens.....</b>	<b>106</b>
<b>12</b>	<b>Quellenverzeichnis .....</b>	<b>110</b>

## 1 Ausgangslage und Aufgabenstellung

Die Fahrleistung von Kraftfahrzeugen (Kfz), also die Summe der Kilometer, die von Kfz innerhalb eines definierten Zeitraumes auf einem definierten Straßennetz zurückgelegt werden, ist eine zentrale Kenngröße zur Beschreibung der Inanspruchnahme der Verkehrsinfrastruktur und damit der Nachfrage im motorisierten Verkehr. Fahrleistungskennzahlen werden in den unterschiedlichsten Zusammenhängen verwendet. So dient die Fahrleistung unter anderem als Bezugsgröße zur Beurteilung der Verkehrssicherheit sowie als Indikator für verkehrsbedingte Umweltbelastungen. Fahrleistungskennzahlen werden regelmäßig von der Bundesanstalt für Straßenwesen (BASt), dem Kraftfahrt-Bundesamt (KBA) und der Bundesanstalt für Güterverkehr (BAG) sowie vom Deutschen Institut für Wirtschaftsforschung (DIW) veröffentlicht.

Die BASt weist auf der Grundlage von Daten aus bundesweiten Verkehrszählungen Kfz-Fahrleistungen auf Bundesautobahnen und den wichtigsten außerörtlichen Straßenklassen aus. Das KBA bestimmt im Rahmen der amtlichen Güterkraftverkehrsstatistik laufend die Beförderungsleistung des Straßengüterverkehrs (inländische Lkw) aufgrund von stichprobenartigen Erhebungen mit Hilfe von Fahrtenprotokollen, aus denen auch die im Rahmen der Güterbeförderung erbrachte Fahrleistung der betreffenden Kfz-Gruppe hochgerechnet werden kann. Die von BASt und KBA publizierten Fahrleistungskennzahlen sind das Resultat der Hochrechnung laufend anfallender empirischer Daten über Fahrleistungen (Fahrleistungen auf Streckenabschnitten des Netzes bzw. Fahrleistungen von in Deutschland zugelassenen Kraftfahrzeugen) mit Hilfe einschlägiger Verfahren der Stichprobentheorie.

Die vom DIW jährlich veröffentlichten Fahrleistungskennzahlen sind demgegenüber das Ergebnis von Modellrechnungen, die nicht auf laufend erhobenen Fahrleistungsdaten beruhen. Eingangsgrößen sind hier vielmehr im Wesentlichen Informationen über den Kraftstoffabsatz und den spezifische Kraftstoffverbrauch von Fahrzeugen, aus methodischer Sicht handelt es sich also um Fahrleistungsschätzungen auf der Basis von Hilfsmerkmalen („Kraftstoffverbrauchsrückrechnung“). Wie bei den KBA-Hochrechnungen fließen auch in das DIW-Schätzverfahren Kfz-Bestandszahlen als weitere Information ein. Da die DIW-Ergebnisse jährlich in „Verkehr in Zahlen“ veröffentlicht werden, haben die betreffenden Kennzahlen halbamtlichen Charakter und finden entsprechend große Aufmerksamkeit und weite Verbreitung.

Es besteht allgemein Einigkeit darüber, dass die Fahrleistungen der in Deutschland zugelassenen Fahrzeuge am genauesten mit Hilfe von speziellen Repräsentativerhebungen in Form von Halterbefragungen ermittelt werden können. Wegen des damit verbundenen Aufwands werden solche Erhebungen allerdings nur in großen Abständen durchgeführt, sie haben also den Charakter von Einzeluntersuchungen („Stand-alone-Studien“). Solche spezielle „Fahrleistungserhebungen“ gab es in Deutschland zuletzt 1993 und 2002.

Die grundlegende Bedeutung spezieller Fahrleistungserhebungen für die Verkehrsstatistik wird unter anderem auch darin sichtbar, dass vor dem Hintergrund der Ergebnisse der Fahrleistungserhebung 2002 und weiterer zeitlich paralleler Verkehrserhebungen (MiD 2002, KiD 2002) die DIW-Modellrechnungen revidiert worden sind. Diese Revidierung war erforderlich geworden, weil - nicht nur im Verkehrsbereich - Modellrechnungen, die sich ausschließlich auf Hilfsmerkmale stützen, dazu tendieren, im Zeitablauf immer stärker von den Ergebnissen empirischer Erhebungen abzuweichen. Dabei öffnet sich die Schere in der Regel umso weiter, je länger die letzte zur Kalibrierung des Berechnungsverfahrens verwendete Erhebung inzwischen zurückliegt. Vor diesem Hintergrund ist es erklärlich, dass im Hinblick auf die laufende Fortschreibung von Fahrleistungskennzahlen ein zunehmendes Interesse an statistischen Verfahren besteht, die auf aktuellen empirischen Daten zur Kfz-Fahrleistung beruhen.

Da die Fortschreibung von Eckwerten der Fahrleistung eine Schätzung der Veränderung der Fahrleistung im Zeitverlauf voraussetzt, bietet sich aus methodischer Sicht in erster Linie die Etablierung eines Fahrleistungspanels als Grundlage einer laufenden Fahrleistungsschätzung an. Mit Paneldaten (hier: wiederholte Beobachtung derselben Fahrzeuge im Hinblick auf das Merkmal Fahrleistung) kann man nämlich zeitliche Veränderungen einer bestimmten Zielgröße (hier: Fahrleistung) mit höherer Genauigkeit schätzen als auf der Grundlage einer Serie von unabhängigen Stichproben gleichen Umfangs. Genauigkeit hat ihren Preis: Die Kosten der Etablierung und Pflege eines Panels sind nicht unbeträchtlich, wie man aus vorliegenden Erfahrungen mit dem Deutschen Mobilitätspanel weiß.

Eine kostengünstige Alternative zu einem speziellen Fahrleistungspanel könnten unter Umständen die Daten sein, die in großer Zahl bei den Hauptuntersuchungen (HU) der Technischen Überwachungsinstitutionen wie z.B. TÜV und Dekra verfügbar sind. Dort wird nämlich für jedes zur Hauptuntersuchung vorgeführte Fahrzeug im abschließenden Prüfbericht unter anderem auch der Kilometerstand am Tag der Hauptuntersuchung festgehalten. Eine ganze Reihe von Prüforganisationen hat inzwischen die Bereitschaft bekundet, für Zwecke einer Fahrleistungsstatistik Daten aus Hauptuntersuchungen bereitzustellen. Die Nutzung von HU-Daten, die als „Prozessdaten“ ja ohnehin anfallen, hat also offenkundig den weiteren Vorzug, dass die Datengewinnung für eine neue Fahrleistungsstatistik keine Beteiligung der Halter der Fahrzeuge voraussetzt; die entsprechende neue Statistik würde also nicht zu irgendwelchen weiteren Belastungen der Kfz-Halter als auskunftgebende Einheiten oder gar als Berichtspflichtige führen.

Bei allen genannten positiven Eigenschaften der HU-Daten stellt sich gleichwohl die Frage nach der methodischen und technisch-organisatorischen Machbarkeit einer auf Prozessdaten der Prüforganisationen beruhenden Fahrleistungsstatistik. Da das Hauptziel einer laufenden Fahrleistungsstatistik die Ermittlung der „aktuellen“ Kfz-Fahrleistungen ist, insbesondere der Fahrleistungen im letzten abgelaufenen Kalenderjahr, ergeben sich aus methodischer Sicht bei Verwendung

von HU-Daten zur Schätzung von periodenbezogenen Fahrleistungen die drei folgenden Kernprobleme:

- Aus den Tachoständen von Fahrzeugen bei der Hauptuntersuchung, d.h. aus Angaben zur gesamten bisherigen Fahrleistung der in einem bestimmten Prüfzeitraum vorgeführten Fahrzeuge, muss der Totalwert der Fahrleistung in einer bestimmten Periode (Kalenderjahr) geschätzt werden. Da die bei der Hauptuntersuchung in Form des Tachostandes festgestellte „bisherige Lebenszeitfahrleistung“ eines Fahrzeugs bei den meisten Fahrzeugen ganz überwiegend vor der Periode erbracht worden ist, für die das Fahrleistungstotal geschätzt werden soll, erscheint eine Nutzung von schwerpunktmäßig „historischen“ HU-Daten zur Ermittlung von Fahrleistungen für Zeitperioden „am aktuellen Rand“ problematisch (Problem der *Validität der HU-Daten*).
- Die HU-Daten liefern naturgemäß Fahrleistungsinformationen nur für diejenigen Fahrzeuge, die im fraglichen Bezugszeitraum (Prüfjahr) an einer Hauptuntersuchung teilgenommen haben. Hier stellt sich unmittelbar die Frage, ob die im Bezugszeitraum geprüften Fahrzeuge als Zufallsstichprobe aus der Gesamtheit aller im betreffenden Zeitraum zugelassenen Fahrzeuge angesehen werden können (Problem der *Repräsentativität der geprüften Fahrzeuge*).
- Da die Bereitschaft zur Datenlieferung nicht bei allen Prüforganisationen vorausgesetzt werden kann, dürften die für eine neue Fahrleistungsstatistik tatsächlich verfügbaren HU-Daten eines bestimmten Bezugszeitraums in aller Regel nur eine Teilmenge der betreffenden HU-Daten insgesamt darstellen. Ob diese Teilmenge als Zufallsstichprobe aus der Gesamtheit aller Kfz-Hauptuntersuchungen im betrachteten Zeitraum angesehen werden kann, ist nicht sicher. Verzerrungen könnten z.B. dann entstehen, wenn es einen Zusammenhang gibt zwischen der Fahrleistung eines Fahrzeugs und der Wahlentscheidung des Fahrzeughalters im Hinblick auf die Prüforganisation, bei welcher das Fahrzeug vorgeführt wird (Problem der *Repräsentativität der tatsächlich bereitgestellten HU-Stichprobe*).

Ein Vorläuferprojekt des BMVBS („Hochrechnung der Kfz-Fahrleistung auf der Basis der Ablesungen der Wegstreckenzähler bei Kfz-Hauptuntersuchungen“, Projekt-Nr. 70.0800/2006/) kam zu dem Ergebnis, dass eine Schätzung jahresbezogener Kfz-Fahrleistungen auf der Basis von HU-Daten grundsätzlich machbar ist und die hierfür benötigten Daten verfügbar sind.

Mit dem vorliegenden Forschungsprojekt (FE-Nr. 70.0832/2008) werden die Bemühungen um die Schaffung einer neuen, auf Prozessdaten aus Hauptuntersuchungen beruhenden Fahrleistungsstatistik fortgeführt. Kernaufgabe des Projekts ist die Entwicklung eines statistischen Verfahrens zur Hochrechnung kalenderjahrbezogener Kfz-Fahrleistungen auf der Basis von Kilometerstands-



daten aus Kfz-Hauptuntersuchungen und die programmtechnische Umsetzung dieses Verfahrens mit Hilfe der Statistik-Standardsoftware SPSS.

Bei aller Umsetzungs- und Anwendungsorientierung kommt es entscheidend darauf an, sich von den reinen Datenfragen einer möglichen auf HU-Daten basierenden Fahrleistungsstatistik zu lösen und zum *statistisch-methodischen Kern* des hier vorliegenden Problems vorzustoßen. Im Folgenden wird deshalb zunächst ein allgemeines statistisches Rahmenkonzept entwickelt, in welches das hier zu erarbeitende Hochrechnungsverfahren methodisch eingeordnet werden kann.

## 2 Methodisches Rahmenkonzept der Untersuchung

### 2.1 Fahrleistung als Gegenstand der empirischen Forschung

Der Begriff „Fahrleistung“ hat in der Umgangssprache aber auch in der Wissenschaft verschiedene Begriffsinhalte. So versteht man in der Fahrzeugtechnik unter Fahrleistung die technische Leistungsfähigkeit eines Kraftfahrzeugs, die in Merkmalen wie Höchstgeschwindigkeit, Beschleunigungsvermögen und maximale Steigfähigkeit zum Ausdruck kommt. Auch in der Verkehrswissenschaft und der Verkehrsstatistik wird der Begriff Fahrleistung - gemeint ist hier die Fahrleistung von Kraftfahrzeugen (nichtspurgeführte Straßenfahrzeuge<sup>1</sup>) - aus unterschiedlichen Perspektiven betrachtet.

#### 2.1.1 Grundlegende Definitionen und Konzepte

In einem 1978 veröffentlichten Projektgruppenbericht der Bundesanstalt für Straßenwesen (BASt)<sup>2</sup> findet sich folgende allgemeine Begriffsdefinition:

„Fahrleistungen sind die Summen aller im Straßenverkehr je Zeiteinheit von den Fahrern der Fahrzeuge oder von den Fahrzeugen selbst zurückgelegten Wegstrecken“.

Wenn speziell die von Fahrzeugen zurückgelegte Wegstrecke gemeint ist, so wird im genannten Projektgruppenbericht von „fahrzeugbezogener Fahrleistung“ gesprochen, deren Dimension dort (ohne nähere Erläuterung) mit  $[Fz \cdot L / T]$  angegeben wird. Bei dieser Betrachtungsweise stellen also Fahrzeuge die Untersuchungseinheiten dar, die über einen gewissen Untersuchungszeitraum hinweg betrachtet werden. Unter der Fahrleistung eines Fahrzeugs versteht man dann die vom betreffenden Fahrzeug im Untersuchungszeitraum insgesamt zurückgelegte Wegstrecke.

Gegenstand der hier vorliegenden Untersuchung ist die fahrzeugbezogene Fahrleistung im Sinne der obigen Definition.

---

<sup>1</sup> Bei Schienenfahrzeugen wird von Betriebsleistung gesprochen.

<sup>2</sup> Grundlegende Darstellungen zum Themenkomplex Fahrleistungen findet man in BASt (Hrsg.): Fahrleistungen im Straßenverkehr - Datenbedarf, Erhebungsmethoden und Realisierungsempfehlungen zu Fahr- und Verkehrsleistungsstatistiken. Projektgruppenberichte der BASt, Bereich Straßenverkehrstechnik, Köln, Dezember 1978

Eine detaillierte Darstellung der konzeptionellen Grundlagen des verkehrstatistischen Fahrleistungsbegriffs - ausgehend von der Mikroebene der Fahrzeugbewegungen in Raum und Zeit - findet sich im Schlussbericht zur Fahrleistungserhebung 1990 (Hautzinger u.a., 1994, S. 11-16). Dort wird das Konzept der „Fahrzeugtage“ als den Untersuchungseinheiten einer Fahrleistungserhebung entwickelt: Bei einem Kollektiv von N Fahrzeugen, einem Untersuchungsgebiet (Straßennetz) bestimmter Abgrenzung und Größe und einem Untersuchungszeitraum der Länge T (im Fall eines Kalenderjahres ist  $T=365$  [Tage]) besteht die Grundgesamtheit aus  $N \cdot T$  Fahrzeug-Tag-Kombinationen. Die im Untersuchungsgebiet von einem bestimmten Fahrzeug  $i$  während eines bestimmten Tages  $t$  zurückgelegte Wegstrecke  $y_{it}$  - gemessen in der Einheit km/Tag - ist dann ein Merkmal des Fahrzeugtags  $(i, t)$ .

Vor diesem Hintergrund ist unter der Fahrleistung des  $i$ -ten Fahrzeugs im Untersuchungsgebiet während des Untersuchungszeitraums die Summe der tagesspezifischen Wegstrecken

$$(2-1) \quad y_i = y_{i1} + \dots + y_{iT} \quad (i=1, \dots, N)$$

zu verstehen. Ziel einer als Einzeluntersuchung durchgeführten Fahrleistungserhebung ist die Ermittlung des Totalwertes

$$(2-2) \quad Y = y_1 + \dots + y_N$$

der Fahrleistungen der zum betrachteten Kollektiv gehörenden Fahrzeuge im Untersuchungsgebiet während des Untersuchungszeitraums. Bei einer laufenden Fahrleistungsstatistik sind Fahrleistungstotalwerte für aufeinanderfolgende Untersuchungszeiträume (insbesondere Kalenderjahre) zu ermitteln.

### 2.1.2 Spezielle Fahrleistungskennzahlen

Wie bei Hautzinger u.a. (1994) gezeigt wird, lassen sich aus diesem allgemeinen statistischen Konzept verschiedene spezielle Fahrleistungsdefinitionen ableiten. Allgemein geläufig sind die folgenden Begriffe:

- Wenn es keine Einschränkung im Hinblick auf das Untersuchungsgebiet gibt, aber das Fahrzeugkollektiv nur die im Inland zugelassenen Fahrzeuge umfasst, so wird der Totalwert  $Y$  als „Inländerfahrleistung“ bezeichnet.
- Wenn es keine Einschränkung im Hinblick auf das Fahrzeugkollektiv gibt, aber das Untersuchungsgebiet nur das Inland umfasst (und dementsprechend von den täglichen Weg-

strecken  $y_{it}$  nur der im Inland zurückgelegte Teil berücksichtigt wird), so wird der Totalwert  $Y$  als „Inlandsfahrleistung“ bezeichnet.

Bezieht man den Fahrleistungstotalwert  $Y$  auf bestimmte Kennzahlen des Fahrzeugbestands (z.B. jahresdurchschnittlicher Fahrzeugbestand, Fahrzeugbestand zur Jahresmitte) oder auf die Gesamtzahl der Fahrzeugtage des Untersuchungszeitraums, so entstehen Verhältniszahlen des Typs

- mittlere Jahresfahrleistung (km pro Fahrzeug und Jahr)
- mittlere Tagesfahrleistung (km pro Fahrzeug und Tag).

In den bereits zitierten Arbeiten finden sich Typologien der Merkmale, welche zur Aufgliederung des Fahrleistungstotalwerts in Frage kommen. Denkbar sind Aufgliederungen nach

- Merkmalen des Fahrzeugs
- Merkmalen des Fahrzeughalters
- Merkmalen des Fahrbetriebs, d.h. der einzelnen Fahrzeugfahrten, aus welchen der Fahrleistungstotalwert resultiert (fahrerbezogene, ladungsbezogene, straßenseitige, regionale, zeitliche Aufgliederungsmerkmale).

Für die Ermittlung von Fahrleistungskennzahlen (Fahrleistungstotalwerte und daraus abgeleitete weitere statistische Kennzahlen zur Fahrleistung von Kraffahrzeugen) stehen verschiedene Erhebungs- und Berechnungsmethoden zur Verfügung.

## **2.2 Methoden zur Gewinnung von statistischen Informationen über Fahrleistungen**

Für die Gewinnung von statistischen Informationen über Fahrleistungen gibt es wie in vielen anderen Bereichen der empirischen Forschung prinzipiell zwei Wege (BASt, 1978, S. 30ff):

- Direkte Methode: Erhebung von Fahrleistungsdaten mittels Befragung oder Beobachtung
- Indirekte Methode: Verwendung von Hilfsvariablen zur Abschätzung der Fahrleistung

Während die indirekte Methode ihrer Natur nach ein sekundärstatistisches Konzept ist, kommen bei der direkten Methode sowohl primär- als auch sekundärstatistische Ansätze in Frage.

## 2.2.1 Primärstatistische Ansätze

### Befragung

Im Rahmen von Befragungen zur Fahrleistung von Kraftfahrzeugen kommen als Erhebungseinheiten in erster Linie natürlich Kfz-Halter in Betracht. Soweit allerdings kein Fahrzeug- oder Halterregister als Auswahlgrundlage zur Verfügung steht, können Fahrleistungsdaten aber auch durch Haushalts- oder Unternehmensbefragungen gewonnen werden. Bei Befragungen sind verschiedene Ansätze zur Erhebung von Einzeldaten über fahrzeugbezogene Fahrleistungen gebräuchlich:

- retrospektive Befragung zur Fahrleistung in einem zurückliegenden Berichtszeitraum (Länge meist 1 Jahr),
- Erfassung der einzelnen Fahrten eines Fahrzeugs an einem oder mehreren Stichtagen und Ermittlung der Tagesfahrleistung als Summe der Einzelfahrlängen (z.B. MID 2008; Mobilitätspanel: Erhebung zur Alltagsmobilität; KiD 2010; amtliche Güterkraftverkehrsstatistik),
- wiederholte Abfrage des Tachostands an mindestens 2 aufeinanderfolgenden Stichtagen (z.B. Fahrleistungserhebung 2002; Mobilitätspanel: Erhebung von Fahrleistung und Kraftstoffverbrauch).

### Beobachtung

Als Grundform der Beobachtung zur Gewinnung von Fahrleistungsdaten kann die Verkehrsbeobachtung (Verkehrszählung) am Ort des Verkehrsgeschehens, d.h. im Straßennetz betrachtet werden. Dabei werden die Straßen des Erhebungsgebiets in Zählabschnitte unterteilt, die so abzugrenzen sind, dass sich in ihnen durch zu- und abfließende Verkehrsmengen nur geringfügige Veränderungen der Verkehrsstärken ergeben können. Jeder beobachteten Fahrzeugbewegung innerhalb eines Zählabschnitts wird dann die zurückgelegte Wegstrecke zugeordnet, die sich nach der Länge des Zählabschnitts bemisst. Multipliziert man die in den einzelnen Zählabschnitten im Bezugszeitraum gezählten Fahrzeugbewegungen mit der Länge der jeweiligen Zählabschnitte, und summiert man schließlich diese Produktwerte über alle Zählabschnitte, so erhält man im Prinzip den Totalwert der Fahrleistungen im Bezugszeitraum nach dem Inlandskonzept (BASt, 1978, S. 56ff). Für das skizzierte Beobachtungsverfahren, das in der Praxis als Stichprobenverfahren mit zufällig ausgewählten Zählabschnitten und Beobachtungszeitintervallen durchgeführt werden sollte, wurde aktuell ein mathematisch-statistisch fundiertes Verfahren zur Hochrechnung und Varianzschätzung entwickelt. Vgl. Hautzinger, Pfeiffer und Schmidt (2010).

Die bisher betrachteten Methoden stellen Primärerhebungen dar. Daneben kommen aber auch sekundärstatistische Ansätze zur Gewinnung von Informationen über Fahrleistungen in Betracht.

## **2.2.2 Sekundärstatistische Ansätze**

### **Hochrechnung von Prozessdaten mit Fahrleistungsbezug**

Direkte Methoden zur Ermittlung von Fahrleistungskennzahlen setzen nicht zwangsläufig spezielle Fahrleistungserhebungen voraus. Vielmehr können die zur Schätzung der Fahrleistung auszuwertenden Daten schon in einem anderen Zusammenhang erhoben worden sein. Im Hinblick auf Fahrleistungen sind hier insbesondere

- Daten des Lkw-Mautsystems und
- Daten aus Kfz-Hauptuntersuchungen

anzusprechen, die beim Betreiber Toll Collect bzw. den Technischen Überwachungsorganisationen im normalen Geschäftsgang anfallen („Prozessdaten“). Die Verwendung von Daten aus Kfz-Hauptuntersuchungen (HU) wird im schon mehrfach zitierten BASt-Projektgruppenbericht unter der Überschrift „Mitwirkung von bestimmten Dienstleistungsinstitutionen bei der Erhebung von Fahrleistungen im Straßenverkehr“ ausführlich diskutiert (BASt, 1978, S.62-68).

Maut- wie auch HU-Daten sind aus statistischer Sicht das Ergebnis von Beobachtungen. Im ersten Fall handelt es sich um eine apparative Beobachtung im fließenden Verkehr (automatische Erfassung von Lkw-Vorbeifahrten an Mautkontrollbrücken im BAB-Netz). Im zweiten Fall besteht die (nichtapparative) Beobachtung im Ablesen des vom Wegstreckenzähler angezeigten Kilometerstandes. Während auf der Basis von Mautdaten Fahrleistungen nach dem Inlandskonzept ermittelt werden können, resultieren aus HU-Daten Fahrleistungen nach dem Inländerkonzept.

### **Fahrleistungsberechnung auf der Basis aggregierter Kraftstoffverbrauchswerte**

Dieses indirekte Verfahren unterscheidet sich von den übrigen Methoden fundamental dadurch, dass zur Ermittlung des Fahrleistungstotalwerts keine Daten über fahrzeugbezogene Fahrleistungen verwendet werden. Stattdessen wird die Tatsache ausgenutzt, für ein bestimmtes Fahrzeugkollektiv ein Zusammenhang zwischen der Hilfsvariablen „Kraftstoffverbrauch im Untersuchungszeitraum“ und der interessierenden Zielgröße „Fahrleistung im Untersuchungszeitraum“ besteht. Die Hilfsvariable wird dabei nicht speziell erhoben, sondern vielmehr in Aggregatdatenform aus sekundärstatistischen Quellen (Mineralölwirtschaft) übernommen.

Die Grundgleichung des Berechnungsverfahrens lautet:

$$(2-3) \quad Y = X/\mu_x$$

wobei Y die Gesamtfahrleistung (in km) des Fahrzeugkollektivs im Untersuchungszeitraum, X den Gesamtkraftstoffverbrauch (in Liter) des Fahrzeugkollektivs im Untersuchungszeitraum und  $\mu_x$  den durchschnittlichen Kraftstoffverbrauch des Fahrzeugkollektivs (in Liter/km) bezeichnet. Der Proportionalitätsfaktor  $1/\mu_x$  wird dabei nicht regressionsanalytisch geschätzt (dazu wären ja Informationen über die abhängige Variable Y erforderlich), sondern aus Herstellerangaben und Testberichten zum spezifischen Kraftstoffverbrauch von Fahrzeugen übernommen.

Das Verfahren ist seiner Natur nach weniger zur originären Ermittlung der Fahrleistungen sondern eher zur kurzfristigen Fortschreibung von auf andere Weise ermittelten Fahrleistungsgrößen geeignet (BASt, 1978, S. 69-72). Eine Darstellung des aktuellen Stands der laufenden Verfeinerung und Anpassung der methodischen Grundlagen des oben skizzierten Verfahrens der „Kraftstoffverbrauchsrückrechnung“ in der Bundesrepublik Deutschland findet sich in der Arbeit von Kalinowska, Kloas, Kuhfeld und Kunert (2005).

### **2.2.3 Eignung der verschiedenen Verfahren zur Fahrleistungsermittlung**

Die Vielzahl und methodische Vielfalt der möglichen Ansätze zur Ermittlung von Fahrleistungen im Straßenverkehr führt zu der Frage, in welchem Verhältnis die verschiedenen Methoden zueinander stehen. Der BASt-Projektgruppenbericht diskutiert in diesem Zusammenhang ausführlich die „konkurrierenden“ und „komplementären“ Erhebungs- bzw. Berechnungsverfahren, wobei Letztere dadurch gekennzeichnet sind, dass zur Erreichung des Erhebungsziels ihr kombinierter Einsatz vonnöten ist (BASt, 1978, S. 73-100).

Wenn es um die Ermittlung von Eckwerten für die Fahrleistungsstatistik geht, so ist ganz eindeutig die Befragung von Kfz-Haltern - in Deutschland zuletzt als „Fahrleistungserhebung 2002“ durchgeführt – das zu präferierende Erhebungsmodell (vgl. Hautzinger u.a., 2005). Soll neben der Inländer- auch die Inlandsfahrleistung bestimmt werden, so ist die Kfz-Halterbefragung mit Verkehrszählungen zu kombinieren (vgl. Hautzinger, Stock und Schmidt, 2005).

Das Anwendungsgebiet der Methoden, die auf sekundärstatistischen Individual- oder Aggregatdaten beruhen, ist nicht die Ermittlung von Eckwerten der Fahrleistung, sondern die Fortschreibung von Fahrleistungskennzahlen. Die hauptsächlichen Gründe für die mangelnde Eignung der sekundärstatistischen Ansätze als Verfahren zur Eckwertbestimmung sind (BASt, 1978, S. 76):

- unscharfe Abgrenzungen der Fahrzeugkollektive, auf welche sich die Fahrleistungskennzahlen beziehen,
- unvollständige Erfassung der für eine Fahrleistungsermittlung relevanten Sachverhalte und
- unvermeidliche Einbeziehung von Arbeitshypothesen, mit denen Informationslücken geschlossen werden sollen.

Die genannten Schwachpunkte der sekundäranalytischen Verfahren gelten im Übrigen nicht nur im Zusammenhang mit der Fahrleistungsstatistik, sie treffen vielmehr für nahezu alle Bereiche der Wirtschafts- und Sozialstatistik zu. „Die zunehmende Automatisierung von Karteiführung und Verwaltungsabläufen ... mag zum Vordringen von Sekundärstatistiken ... führen. Sie können jedoch die wesentlich flexibleren Primärerhebungen nicht ersetzen. ... Auch die vermehrte Durchführung von Schätzungen (u.a. mit Unterstützung von Datenbanken) kann aufwendige Erhebungen nur z. T. entbehrlich machen.“ (von der Lippe, 1990). Gleichwohl sind sekundärstatistische Daten wertvolle Bestandteile von statistischen Informationssystemen.

Vor diesem Hintergrund wird im Folgenden untersucht, durch welche statistisch-methodischen Ansätze das Potenzial von Prozessdaten der Technischen Überwachungsinstitutionen (Tachostandsablesung bei der Kfz-Hauptuntersuchung) für Zwecke der Fahrleistungsstatistik bestmöglich erschlossen werden kann.



### 3 Statistisches Grundmodell der Fahrleistungen von Kraftfahrzeugen

Für die Entwicklung eines statistischen Verfahrens zur Schätzung von *zeitraumbezogenen* Fahrleistungen auf der Grundlage von *zeitpunktbezogenen* Tachostandsdaten wird ein konzeptueller und begrifflicher Rahmen benötigt. Dieser Rahmen wird im vorliegenden Kapitel bereitgestellt. Das hier beschriebene statistische Modell ist ein *Modell für die Grundgesamtheit*, es dient vorrangig dazu, diejenigen Fahrleistungskennzahlen präzise zu charakterisieren, um deren Schätzung es im Weiteren geht. Aus dem Fahrleistungsmodell für die Grundgesamtheit der Fahrzeuge können aber auch bereits wichtige Folgerungen für die Struktur des zu entwickelnden Schätzverfahrens abgeleitet werden.

Im Folgenden wird von der Vorstellung ausgegangen, dass Fahrzeuge und deren Fahrleistungen im Zeitverlauf beobachtet werden und dass dabei die Kalenderzeit in Perioden (Zeitabschnitte gleicher Länge, z.B. Jahre) untergliedert ist. Das nachfolgend vorgestellte allgemeine statistische Modell der Fahrleistungen von Kraftfahrzeugen bildet den theoretischen Hintergrund für die Ausarbeitung eines Verfahrens zur Schätzung von periodenbezogenen Fahrleistungen auf der Grundlage von Daten zur Fahrleistung vom Tag der Erstanmeldung bis zu einem bestimmten Stichtag in der betrachteten Periode (Tag der Hauptuntersuchung). Obschon das spätere Hochrechnungsverfahren auf Monatsbasis formuliert wird, beruht das Modell für die Grundgesamtheit auf dem Konzept der „Fahrzeugtage“ gemäß Abschnitt 2.2.1.

Die Modellierung der Fahrleistungen berücksichtigt explizit den Anmeldestatus der Fahrzeuge an den einzelnen Tagen der betrachteten Periode. Die Berücksichtigung des Anmeldestatus im Modell ist unter anderem deshalb wichtig, weil der Hochrechnungsrahmen durch Kfz-Bestandsdaten des Zentralen Fahrzeugregisters (ZFZR) gebildet wird und diesem Register seit 2008 nur noch die Zahl der zu einem bestimmten Stichtag angemeldeten Fahrzeuge entnommen werden kann (zwar enthält das Register auch außerbetriebgesetzte Fahrzeuge, doch ist der Grund der Außerbetriebsetzung wie z.B. Abmeldung oder vorübergehende Stilllegung nicht mehr nachvollziehbar).

#### 3.1 Periodenbezogene Fahrleistung von Einzelfahrzeugen

Wie oben beschrieben, stellt bei Analysen auf der Mikroebene jede Kombination  $(i, t)$  von Fahrzeug und Kalendertag, kurz jeder „Fahrzeugtag“, eine Untersuchungseinheit dar. Untersucht man aus einem gewissen Fahrzeugkollektiv des Umfangs  $N$  ein einzelnes Fahrzeug  $i$  an den verschiedenen Tagen  $t$  einer bestimmten Periode (Untersuchungszeitraum) der Länge  $T$ , so ist vorrangig

der Anmeldestatus des Fahrzeugs und – soweit am betreffenden Tag angemeldet – die Fahrleistung des Fahrzeugs von Interesse. Hierzu wird folgende Notation eingeführt ( $i=1,\dots,N$ ;  $t=1,\dots,T$ ):

$a_{it}$  Indikatorvariable für den Anmeldestatus<sup>3</sup> des  $i$ -ten Fahrzeugs am Tag  $t$  des Untersuchungszeitraums (1 = angemeldet, 0 = nicht angemeldet)

$y_{it}$  Fahrleistung des  $i$ -ten Fahrzeugs am Tag  $t$  des Untersuchungszeitraums gemessen in km

Im Folgenden wird in leichter Idealisierung unterstellt, dass Fahrzeuge nur an Tagen, an denen sie angemeldet sind, Fahrleistungen erbringen, d.h.  $y_{it} = 0$  falls  $a_{it} = 0$  und  $y_{it} \geq 0$  falls  $a_{it} = 1$ .

Geht man von dieser Vorstellung aus, so interessiert zunächst die Gesamtzahl aller Fahrzeugtage, an denen beim Fahrzeug  $i$  im Untersuchungszeitraum überhaupt Fahrleistungen entstehen können. Diese Gesamtzahl ist durch

$$(3-1) \quad a_i = a_{i1} + a_{i2} + \dots + a_{iT}$$

gegeben und wird im Folgenden *Anmeldedauer des  $i$ -ten Fahrzeugs im Untersuchungszeitraum* genannt (Anmeldedauer in Tagen). Zum betrachteten Kollektiv gehören nur Fahrzeuge, die an mindestens einem Tag des Untersuchungszeitraums angemeldet sind, es gilt also  $1 \leq a_i \leq T$ .

Die durch

$$(3-2) \quad y_i = y_{i1} + y_{i2} + \dots + y_{iT}$$

gegebene *Fahrleistung des  $i$ -ten Fahrzeugs im Untersuchungszeitraum* (in km) wird im Folgenden kurz *periodenbezogene Fahrleistung* genannt. Die periodenbezogene Fahrleistung (3-2) des  $i$ -ten Fahrzeugs kann unter Verwendung der Indikatorvariablen  $a_{it}$  für den Anmeldestatus auch wie folgt geschrieben werden:

$$(3-3) \quad y_i = y_{i1} \cdot a_{i1} + \dots + y_{iT} \cdot a_{iT} .$$

---

<sup>3</sup> Dem Zentralen Fahrzeugregister (ZFZR) kann nur die Zahl der jeweils angemeldeten Fahrzeuge entnommen werden, nicht aber die Zahl der vorübergehend stillgelegten Fahrzeuge. Da die Fahrzeugbestände des ZFZR den Hochrechnungsrahmen für eine Schätzung der Jahresfahrleistung liefern, wird für jeden Stichtag innerhalb des Untersuchungsjahres jeweils nur der Bestand der angemeldeten Fahrzeuge betrachtet. Ein Fahrzeug, das im Verlauf des Untersuchungsjahres erstmals zugelassen bzw. endgültig stillgelegt wird, hat im vorliegenden Zusammenhang für die Tage vor der Erstzulassung bzw. nach der endgültigen Stilllegung jeweils den Status „nicht angemeldet“.

Bezeichnet man mit  $m_i = y_i/T$  die mittlere Fahrleistung des  $i$ -ten Fahrzeugs pro Tag des Untersuchungszeitraums und mit  $m_i^{(a)} = y_i/a_i$  die mittlere Fahrleistung des  $i$ -ten Fahrzeugs pro Anmelde- tag des Untersuchungszeitraums, so erhält man für die periodenbezogene Fahrleistung die Darstellung

$$(3-4) \quad y_i = T \cdot m_i = a_i \cdot m_i^{(a)}$$

Die Fahrleistung eines einzelnen Fahrzeugs im Untersuchungszeitraum (periodenbezogene Fahrleistung) ist also

- proportional zur Anmeldungsdauer im Untersuchungszeitraum sowie
- proportional zur mittleren Fahrleistung pro Anmelde- tag.

Im Folgenden wird jetzt ein Fahrzeugkollektiv als Ganzes während eines bestimmten Untersuchungszeitraums (z.B. Kalenderjahr) betrachtet. Gegenstand des Interesses ist die Gesamtsumme der Fahrleistungen der zum Kollektiv gehörenden Fahrzeuge im betreffenden Untersuchungszeitraum. Diese Gesamtsumme ist offenkundig eine Fahrleistungskennzahl nach dem Inländerkonzept.

## 3.2 Periodenbezogene Fahrleistung eines Fahrzeugkollektivs

### 3.2.1 Anmeldequote und mittlerer Bestand angemeldeter Fahrzeuge

Für eine quantitative Fahrleistungsanalyse ist es zunächst erforderlich, den Anmeldestatus der Fahrzeuge des betrachteten Kollektivs im Untersuchungszeitraum zahlenmäßig zu beschreiben. Dazu wird folgende Notation eingeführt:

T Länge des Untersuchungszeitraums (T=365 Tage)

N Umfang des betrachteten Fahrzeugkollektivs (Zahl der Fahrzeuge mit mindestens 1 Anmelde- tag im Untersuchungszeitraum)

$A = a_1 + a_2 + \dots + a_N$  Gesamtzahl der Anmelde- tage der N Fahrzeuge im Untersuchungs- zeitraum (Totalwert der Anmelde- dauer in Tagen)

$\mu_A = (a_1 + \dots + a_N)/N$  mittlere Anzahl der Anmelde- tage pro Fahrzeug im Untersuchungs- zeitraum („mittlere Anmelde- dauer“ der Fahrzeuge in Tagen)

$\alpha = A / (N \cdot T)$  Anmeldequote (Verhältnis zwischen Gesamtzahl der Anmelde- tage und Gesamtzahl der Fahrzeug- tage im Untersuchungszeitraum)

$B_t = a_{1t} + a_{2t} + \dots + a_{Nt}$  Gesamtzahl der am Tag  $t$  angemeldeten Fahrzeuge<sup>4</sup> ( $0 \leq B_t \leq N$ ); dieser tagesspezifische Bestand angemeldeter Fahrzeuge entspricht der Zahl der Anmeldetage von Fahrzeugen des betrachteten Kollektivs, welche auf den Kalendertag  $t$  entfallen

$\mu_B = (B_1 + \dots + B_T)/T$  mittlere Anzahl angemeldeter Fahrzeuge pro Kalendertag („mittlerer Bestand angemeldeter Fahrzeuge“; Bestandsmittelwert über alle Tage des Untersuchungszeitraums<sup>5</sup>)

Da die Summe der tagesspezifischen Bestände angemeldeter Fahrzeuge gleich der Gesamtzahl der Anmeldetage im Untersuchungszeitraum ist ( $B_1+B_2+\dots+B_T = A$ ), gilt

$$(3-5) \quad \mu_B \cdot T = \mu_A \cdot N = A .$$

Zwischen der mittleren Anmeldedauer  $\mu_A$  (in Tagen) und dem mittleren Bestand  $\mu_B$  angemeldeter Fahrzeuge besteht gemäß (3-5) also ein Zusammenhang der Form

$$(3-6) \quad \mu_B = (N/T) \cdot \mu_A ,$$

d.h. der mittlere Bestand  $\mu_B$  angemeldeter Fahrzeuge ist gleich der mit dem Faktor  $N/T$  multiplizierten mittleren Anmeldedauer  $\mu_A$  der Fahrzeuge im Untersuchungszeitraum. Nähert sich die mittlere Anmeldedauer dem Wert  $T$  (Länge des Untersuchungszeitraums), so strebt der mittlere Bestand angemeldeter Fahrzeuge gegen  $N$  (Umfang des Fahrzeugkollektivs).

### 3.2.2 Periodenbezogene Fahrleistungskennzahlen

Summiert man die individuellen periodenbezogenen Fahrleistungen (3-2) über alle Fahrzeuge, so erhält man den *periodenbezogenen Fahrleistungstotalwert*

$$(3-7) \quad Y = y_1 + y_2 + \dots + y_N$$

d.h. die Gesamtsumme der im Untersuchungszeitraum vom betreffenden Fahrzeugkollektiv zurückgelegten Kilometer. Dieser Totalwert ist, wie erwähnt, eine auf dem Inländerkonzept beruhenden

---

<sup>4</sup> Wie bereits erwähnt, werden im Zentralen Fahrzeugregister (ZFZR) immer nur die aktuell angemeldeten Fahrzeuge geführt. Die Größe  $B_t$  entspricht demnach dem ZFZR-Bestand am Tag  $t$  des Untersuchungszeitraums.

<sup>5</sup> Wie man sieht, wächst der mittlere Bestand angemeldeter Fahrzeuge mit dem Umfang des Fahrzeugkollektivs ( $N$ ) und der Anmeldequote ( $\alpha$ ).

de Fahrleistungskennzahl, die bei statistischen Veröffentlichungen üblicherweise in der Einheit „Mrd. km“ angegeben wird. Als Periode ist in der Regel ein Kalenderjahr zu betrachten.

Der Fahrleistungstotalwert  $Y$  kann einerseits auf die Gesamtzahl  $A$  der Anmelde tage von Fahrzeugen und andererseits auf die Gesamtzahl  $N \cdot T$  der Fahrzeugtage<sup>6</sup> in der betreffenden Periode bezogen werden. Man erhält so zwei verschiedene Verhältnis zahlen, die als Fahrleistungsmittelwerte interpretiert werden können:

$$(3-8) \quad \mu_Y = Y / A \quad \text{periodenbezogene mittlere Fahrleistung pro Anmelde tag}$$

$$(3-9) \quad \mu^*_Y = Y / (N \cdot T) \quad \text{periodenbezogene mittlere Fahrleistung pro Fahrzeugtag.}$$

Die Kennzahlen (3-8) und (3-9) sind durch die Beziehung

$$(3-10) \quad \mu_Y = \mu^*_Y / \alpha$$

verknüpft, wobei  $\alpha$  die Anmeldequote bezeichnet ( $\alpha = A / (N \cdot T)$ ).

Als weitere periodenbezogene Fahrleistungskennzahl ist die Verhältniszahl

$$(3-11) \quad R = Y / \mu_B$$

zu nennen, d.h. die Gesamtfahrleistung im Untersuchungszeitraum bezogen auf den entsprechenden mittleren Bestand angemeldeter Fahrzeuge (periodenbezogene mittlere Fahrleistung angemeldeter Fahrzeuge)

Im Hinblick auf das zu entwickelnde statistische Verfahren zur Schätzung der Gesamtfahrleistung eines Fahrzeugkollektivs während eines Untersuchungszeitraums der Länge  $T$  (Anzahl Tage) sind die aus (3-8) bzw. (3-9) resultierenden Darstellungen des Fahrleistungstotal  $Y$  von Bedeutung:

$$(3-12) \quad Y = \mu_Y \cdot A = \mu_Y \cdot \mu_B \cdot T$$

---

<sup>6</sup> Die Gesamtzahl  $N \cdot T$  der Fahrzeugtage hat insofern hypothetischen Charakter, als in dieser Zahl auch Fahrzeugtage vor der Erstanmeldung (gilt für Fahrzeuge, die im Untersuchungszeitraum erstmals angemeldet werden) und Fahrzeugtage nach der endgültigen Stilllegung (gilt für Fahrzeuge, die im Untersuchungszeitraum endgültig stillgelegt werden) enthalten sind. Für das eigentliche Hochrechnungsverfahren spielt die Größe  $N \cdot T$  aber keine Rolle, da dort die Zahl  $A$  der Anmelde tage eingeht.

und

$$(3-13) \quad Y = \mu^*_Y \cdot N \cdot T = (\mu^*_Y / \alpha) \cdot \mu_B \cdot T .$$

Gleichung (3-12) zeigt, dass für die Schätzung des Fahrleistungstotals  $Y$  ein Schätzwert für die mittlere Fahrleistung  $\mu_Y$  pro Anmeldetag und ein Schätzwert für den mittleren Bestand  $\mu_B$  angemeldeter Fahrzeuge benötigt wird. Aus (3-13) geht hervor, dass anstelle der mittleren Fahrleistung pro Anmeldetag ( $\mu_Y$ ) auch die mittlere Fahrleistung pro Fahrzeugtag ( $\mu^*_Y$ ) geschätzt werden kann; in diesem Fall wird zusätzlich allerdings noch ein Schätzwert für die periodenbezogene Anmeldequote ( $\alpha$ ) benötigt.

### 3.3 Bisherige Fahrleistung: Tachostand von Fahrzeugen

#### 3.3.1 Bisherige Lebensdauer und bisherige Fahrleistung von Fahrzeugen

Bei den obigen Betrachtungen war stets eine bestimmte Periode als Untersuchungszeitraum (etwa das laufende Kalenderjahr) vorgegeben. Gegenstand des Interesses war die Anmeldedauer und die Fahrleistung von Einzelfahrzeugen bzw. Fahrzeugkollektiven in der betrachteten Periode.

Im Folgenden wird stattdessen nun ein bestimmter Zeitpunkt - hier ein Stichtag - vorgegeben, auf welchen sich die Betrachtungen beziehen. Gegenstand des Interesses ist jetzt die *bisherige Lebensdauer* des Fahrzeugs (am Stichtag erreichtes Alter) und die *bisherige Fahrleistung* des Fahrzeugs (am Stichtag erreichter Tachostand). Unter dem Alter eines Fahrzeugs am Stichtag wird dabei die Länge des Zeitraums zwischen der Erstanmeldung des Fahrzeugs und dem Stichtag verstanden; gemessen wird bei dieser modellhaften Betrachtung das Alter in Tagen. Die bisherige Fahrleistung schließt die Fahrleistung am Stichtag selbst mit ein (Tachostand am Ende des Stichtags).

Zur Verknüpfung von zeitraum- und zeitpunktbezogener Fahrleistungsanalyse wird im Folgenden ein bestimmter fester Untersuchungszeitraum (typischerweise ein Kalenderjahr) und innerhalb des Untersuchungszeitraums ein bestimmter Stichtag  $t$  betrachtet. Untersucht wird als Fahrzeugkollektiv wie zuvor die Gesamtheit der Fahrzeuge mit mindestens 1 Anmeldetag im Untersuchungszeitraum.

Bezeichnet man mit  $d_i$  die Länge des Zeitraums (in Tagen) von der Erstzulassung des  $i$ -ten Fahrzeugs bis zum Beginn des Untersuchungszeitraums<sup>7</sup>, so ist

$$(3-14) \quad k_{it} = d_i + t$$

die *bisherige Lebensdauer* des  $i$ -ten Fahrzeugs gemessen in Tagen (Alter des Fahrzeugs am Tag  $t$ ). Die im Zeitraum der Länge  $k_{it}$  erbrachte Fahrleistung  $z_{it}$  (gemessen in km) wird ganz entsprechend *bisherige Fahrleistung* oder kurz *Tachostand am Tag  $t$*  genannt. Die Verhältniszahl

$$(3-15) \quad m_{zit} = z_{it} / k_{it}$$

stellt die in der Einheit km/Tag gemessene *mittlere bisherige Fahrleistung* des  $i$ -ten Fahrzeugs bezogen auf den Stichtag  $t$  dar. Im später entwickelten Schätzverfahren spielt diese Größe eine wichtige Rolle.

Betrachtet man ein Fahrzeugkollektiv zu einem bestimmten Zeitpunkt  $t$  (Stichtag), so variiert sowohl die bisherige Fahrleistung  $z_{it}$  als auch die bisherige Lebensdauer  $k_{it}$  von Fahrzeug zu Fahrzeug. Fahrzeuge aus dem betrachteten Kollektiv, welche am Stichtag hinsichtlich ihrer bisherigen Lebensdauer übereinstimmen, werden im Folgenden eine *Kohorte* genannt (gleicher Tag der Erstzulassung und damit gleiches Alter am Stichtag).

### 3.3.2 Bisherige versus aktuelle Fahrleistung von Fahrzeugen

Es liegt nahe, die während der ersten  $t$  Tage des Untersuchungszeitraums (z.B. laufendes Kalenderjahr) erbrachte Fahrleistung des  $i$ -ten Fahrzeugs, d.h. die Summe

$$(3-16) \quad u_{it} = y_{i1} + \dots + y_{it}$$

als *aktuelle Fahrleistung* des  $i$ -ten Fahrzeugs zu bezeichnen (Fahrleistung im Untersuchungszeitraum bis zum Stichtag  $t$ ). Handelt es sich beim Stichtag um den letzten Tag des Untersuchungszeitraums (d.h.  $t=T$ ), so gilt  $u_{iT} = y_i$ ; in diesem Fall stimmt die aktuelle Fahrleistung mit der periodenbezogenen Fahrleistung überein.

---

<sup>7</sup> Gemeint ist die Lebensdauer des Fahrzeugs bis zum Beginn des betrachteten Untersuchungszeitraums. Diese Lebensdauer entspricht der Länge des Zeitraums, auf welchen sich die in Abschnitt 3.3.2 definierte „frühere“ Fahrleistung  $f_i$  bezieht.

Ganz entsprechend kann die Fahrleistung  $f_i$ , die das Fahrzeug vor dem betrachteten Untersuchungszeitraum erbracht hat, als *frühere Fahrleistung* bezeichnet werden, da der Zeitraum, auf welchen sich die Fahrleistungsangabe bezieht, schon weiter zurückliegt.

Die Summe aus früherer Fahrleistung  $f_i$  und aktueller Fahrleistung  $u_{it}$  entspricht natürlich der bisherigen Fahrleistung bzw. dem Tachostand am Tag  $t$ :

$$(3-17) \quad z_{it} = f_i + u_{it} .$$

Für ein einzelnes Fahrzeug  $i$  und einen bestimmten Stichtag  $t$  innerhalb des Untersuchungszeitraums kann man die beiden folgenden Fahrleistungsmittelwerte unterscheiden:

$$(3-18) \quad m_{Zit} = z_{it} / k_{it} \quad \text{mittlere bisherige Fahrleistung pro Tag}$$

$$(3-19) \quad m_{Uit} = u_{it} / t \quad \text{mittlere aktuelle Fahrleistung pro Tag.}$$

Zwischen der mittleren bisherigen und der mittleren aktuellen Fahrleistung eines Fahrzeugs besteht natürlich ein Zusammenhang. Bezeichnet man mit

$$(3-20) \quad m_{Fi} = f_i / d_i$$

die mittlere frühere Fahrleistung pro Tag (mittlere Tagesfahrleistung vor Beginn des Untersuchungszeitraums), so erhält man unter Berücksichtigung von (3-14) die Darstellung

$$(3-21) \quad m_{Zit} = z_{it} / k_{it} = (f_i + u_{it}) / k_{it} = (m_{Fi} \cdot d_i + m_{Uit} \cdot t) / k_{it}$$

und somit

$$(3-22) \quad m_{Zit} = m_{Fi} \cdot (d_i / k_{it}) + m_{Uit} \cdot (t / k_{it}).$$

Wie Gleichung (3-22) zeigt, ist die mittlere bisherige Fahrleistung  $m_{Zit}$  gleich dem mit den entsprechenden Zeitanteilen gewogenen arithmetischen Mittel der mittleren früheren Fahrleistung  $m_{Fi}$  (vor dem betrachteten Untersuchungszeitraum) und der mittleren aktuellen Fahrleistung  $m_{Uit}$  (im Unter-



suchungszeitraum)<sup>8</sup>. Wie man sieht, wird die mittlere bisherige Fahrleistung eines einzelnen Fahrzeugs umso stärker von der früheren Fahrleistung bestimmt, je älter das Fahrzeug ist.

Im Folgenden wird für ein Fahrzeugkollektiv die mittlere bisherige Fahrleistung (d.h. der mittlere Tachostand) betrachtet. Dabei wird zwischen dem mittleren Tachostand an einem bestimmten Stichtag innerhalb des Untersuchungszeitraums (stichtagsbezogener Mittelwert des Tachostands über die Fahrzeuge) und dem mittleren Tachostand in der betrachteten Periode (Gesamtmittelwert des Tachostands über Fahrzeuge und Tage) unterschieden.

### 3.3.3 Stichtagsbezogener mittlerer Tachostand von Fahrzeugen

Wenn es um den Mittelwert des Tachostands der Fahrzeuge des Kollektivs<sup>9</sup> an einem Stichtag  $t$  geht, so sind für die Mittelwertberechnung auf jeden Fall nur die Fahrzeuge maßgeblich, deren individuelle Lebenszeit<sup>10</sup> den Tag  $t$  überdeckt. Jedes dieser Fahrzeuge (gleichgültig ob am Tag  $t$  angemeldet oder vorübergehend stillgelegt) weist am Tag  $t$  einen bestimmten Tachostand auf. Dieser wird wie zuvor mit  $z_{it}$  bezeichnet. Fahrzeuge aus dem Kollektiv, die bereits vor dem Tag  $t$  endgültig stillgelegt wurden und Fahrzeuge, deren Erstzulassung erst nach dem Tag  $t$  erfolgt, sind bei der Mittelwertbildung nicht zu berücksichtigen, da für sie die Variable  $z_{it}$  nicht definiert ist. Würde man den eben umrissenen Ansatz wählen, so wären unter den zu berücksichtigenden Fahrzeugen sowohl angemeldete als auch vorübergehend stillgelegte Fahrzeuge<sup>11</sup>.

Wegen der bereits mehrfach angesprochenen Bedeutung des Bestands der angemeldeten Fahrzeuge für die Hochrechnung der Tachostandsdaten wird im Folgenden der *mittlere Tachostand der am Tag  $t$  angemeldeten Fahrzeuge* betrachtet, der durch

$$(3-23) \quad \mu_Z(t) = (z_{1t} \cdot a_{1t} + \dots + z_{Nt} \cdot a_{Nt}) / B_t$$

<sup>8</sup> In den folgenden Ausführungen wird es um die Frage gehen, wie gut die Übereinstimmung zwischen  $m_{Zit}$  und  $m_{Uit}$  ist. Aus der Ablesung des Tachostands an einem Stichtag  $t$  kann nur  $m_{Zit}$ , nicht aber  $m_{Uit}$  ermittelt werden. Für eine Schätzung der Fahrleistung im Untersuchungszeitraum werden aber die „aktuellen“ Fahrleistungen benötigt.

<sup>9</sup> Das Fahrzeugkollektiv des Umfangs  $N$  ist keine zeitpunktbezogene Bestandsgesamtheit im üblichen Sinne; zum Kollektiv gehören vielmehr alle Fahrzeuge, die an mindestens einem Tag des betrachteten Untersuchungszeitraums den Status „angemeldet“ besitzen. Die Zugehörigkeit zum Kollektiv definiert sich also über den Anmeldestatus im Untersuchungszeitraum.

<sup>10</sup> Kalenderzeitabschnitt zwischen Tag der Erstzulassung und Tag der endgültigen Stilllegung.

<sup>11</sup> Der für diese Art der Mittelwertbildung maßgebliche Fahrzeugbestand könnte durch Auszählung des Zentralen Fahrzeugregisters nicht ermittelt werden, da dort die vorübergehend stillgelegten Fahrzeuge nicht (mehr) identifiziert werden können.

gegeben ist. Die Tachostandssumme auf der rechten Seite der Gleichung (3-23) umfasst nur die Tachostände der angemeldeten Fahrzeuge, da für die am Tag  $t$  vorübergehend stillgelegten Fahrzeuge  $a_{it}=0$  gilt.

### 3.3.4 Periodenbezogener mittlerer Tachostand von Fahrzeugen

Betrachtet man den Tachostand eines einzelnen Fahrzeugs  $i$  an seinen insgesamt  $a_i$  Anmeldetagen in der interessierenden Periode, so ist durch

$$(3-24) \quad m_{Zi} = (z_{i1} \cdot a_{i1} + \dots + z_{iT} \cdot a_{iT}) / a_i$$

der *periodenbezogene mittlere Tachostand des  $i$ -ten Fahrzeugs* gegeben (Mittelwert über alle Anmeldetage des Fahrzeugs in der Periode).

Summiert man ganz entsprechend für die betreffende Periode die Tachostände über die Anmeldetage aller Fahrzeuge des Kollektivs und bezieht man diese Summe auf die Gesamtzahl der Anmeldetage, so ergibt sich der periodenbezogene mittlere Tachostand des betrachteten Fahrzeugkollektivs. Die Summe der individuellen Tachostände  $z_{it}$  für die insgesamt  $A$  Anmeldetage der Fahrzeuge des betrachteten Kollektivs in der Periode ist durch

$$(3-25) \quad Z = \sum \sum a_{it} \cdot z_{it} = z_{11} \cdot a_{11} + z_{12} \cdot a_{12} + \dots + z_{1T} \cdot a_{1T} + z_{21} \cdot a_{21} + \dots + z_{NT} \cdot a_{NT}$$

gegeben, da nur die Fahrzeugtage  $(i, t)$  mit  $a_{it}=1$  zu der Merkmalssumme  $Z$  beitragen. Bezieht man diese Summe auf die Zahl der Anmeldetage, so erhält man mit

$$(3-26) \quad \mu_Z = Z / A$$

den *periodenbezogenen mittleren Tachostand angemeldeter Fahrzeuge* gemessen in km.

## 3.4 Das Dilemma: Schätzung periodenbezogener Fahrleistungen auf der Basis stichtagsbezogener Tachostandsdaten

### 3.4.1 Periodenbezogene Kennzahlen als Zielgrößen der Verkehrsstatistik

Sieht man von einigen zeitpunktbezogenen Kennzahlen wie Kraftfahrzeugbestand zum Jahresbeginn oder Anlagevermögen eines Verkehrsbereichs zum Jahresende ab, so liefert die amtliche und halbamtliche Verkehrsstatistik im Wesentlichen periodenbezogene - häufig kalenderjahrbezogene -

Kennzahlen. Dies ist insofern auch ganz natürlich, als der Untersuchungsgegenstand „Verkehr“ eine Ereignisgesamtheit darstellt (Gesamtheit von Ortsveränderungen, die in Raum und Zeit stattfinden).

Selbstverständlich sind auch die Zielgrößen einer Fahrleistungsstatistik periodenbezogene Kennzahlen. Gemäß (3-12) kann für ein bestimmtes Fahrzeugkollektiv der periodenbezogene Fahrleistungstotalwert  $Y$  in der Form

$$(3-27) \quad Y = \mu_Y \cdot \mu_B \cdot T$$

dargestellt werden. In dieser Darstellung wird deutlich, dass die zu schätzende Grundgesamtheitskennzahl  $Y$  von folgenden Faktoren bestimmt wird:

- Länge der Periode in Tagen ( $T$ )
- periodenbezogener mittlerer Bestand angemeldeter Fahrzeuge ( $\mu_B$ )
- periodenbezogene mittlere Fahrleistung pro Anmeldetag ( $\mu_Y$ ).

Stellt z.B. das Kalenderjahr 2009 die Bezugsperiode der Hochrechnung dar, so ist  $T=365$  und für die Schätzung des Fahrleistungstotalwerts  $Y_{2009}$  wird

- ein Schätzwert für den mittleren Bestand angemeldeter Fahrzeuge im Jahr 2009 sowie
- ein Schätzwert für die mittlere Fahrleistung pro Fahrzeuganmeldetag im Jahr 2009 benötigt.

Einen Schätzwert für den jahresdurchschnittlichen Fahrzeugbestand 2009 kann man ohne Weiteres durch wiederholte Auszählung des in Form des ZFZR vorliegenden vollständigen Verzeichnisses der zur Grundgesamtheit gehörenden Fahrzeuge gewinnen (nur angemeldete Fahrzeuge).

Die Schätzung der mittleren Fahrleistung im Bezugsjahr 2009 (km pro Fahrzeug und Anmeldetag) kann demgegenüber nur auf der Basis von Stichprobendaten erfolgen. Es liegt nahe, dass es sich dabei idealerweise um *empirische Daten über im Bezugsjahr erbrachte Fahrleistungen* handeln sollte, wie sie z.B. aus Halterbefragungen des in Abschnitt 2.2.1 beschriebenen Typs resultieren. Stichtagsbezogene Daten zum Tachostand von Fahrzeugen erfüllen diese Anforderung nicht bzw. nicht unmittelbar.

### 3.4.2 Tachostandsdaten als empirisches Material der Verkehrsstatistik

Eine Datensammlung, die für eine gewisse Anzahl  $n$  von Fahrzeugen Informationen zum Tachostand an einem bestimmten Stichtag (Prüftermin) innerhalb einer Bezugsperiode enthält, kann als Stichprobe aus der entsprechenden Gesamtheit aller Fahrzeugtage der betreffenden Periode angesehen werden. Geht man vereinfachend davon aus, dass Fahrzeuge an ihren Prüftagen angemeldet sind, so handelt es sich um eine Stichprobe des Umfangs  $n$  aus der Menge der insgesamt  $A$  Fahrzeuganmeldetage der betrachteten Periode. Kernmerkmale der Untersuchungseinheit „Fahrzeuganmeldetag“ sind die Merkmale „Datum“ und „Kilometerstand“.

Wofür eignen sich nun derartige Daten?

Empirische Tachostandsdaten, bei denen die Stichtage (Tag der Ablesung des Tachostands) über die gesamte Bezugsperiode verteilt sind, sind ihrem Wesen nach in erster Linie dafür geeignet, den durch (3-26) gegebenen mittleren Tachostand  $\mu_z$  angemeldeter Fahrzeuge in der Bezugsperiode zu schätzen. Dies gilt auch für die hier in Rede stehenden HU-Daten, da die Prüftermine der im Verlauf eines Kalenderjahres bei den Prüforganisationen vorgeführten Fahrzeuge über das ganze Kalenderjahr streuen. Wenn es sich bei der Stichprobe der  $n$  „Fahrzeugprüftage“ um eine einfache Zufallsstichprobe aus der Gesamtheit der  $A$  Fahrzeuganmeldetage handeln würde<sup>12</sup>, so wäre für die betreffende Periode (z.B. Kalenderjahr 2009) der mittlere Tachostand  $m_z$  der geprüften Fahrzeuge eine erwartungstreue Schätzung für den mittleren Tachostand  $\mu_z$  angemeldeter Fahrzeuge.

Bei realistischer Betrachtung muss man hier allerdings feststellen, dass Schätzungen des mittleren Tachostands der im betreffenden Kalenderjahr angemeldeten Fahrzeuge im Rahmen der amtlichen oder halbamtlichen Verkehrsstatistik praktisch nicht interessieren. Wenn überhaupt, so könnten vielleicht Fahrzeughersteller im Rahmen ihrer Marktbeobachtung ein gewisses Interesse an derartigen Kennzahlen haben.

---

<sup>12</sup> In diesem Fall hätten alle Fahrzeuganmeldetage des betrachteten Fahrzeugkollektivs in der betrachteten Periode dieselbe Auswahlchance und es könnte prinzipiell ein bestimmtes Fahrzeug auch mit zwei oder mehr Prüftagen in der Stichprobe vertreten sein. Tatsächlich handelt es sich bei den HU-Daten aber nicht um eine solche einfache Zufallsstichprobe; dies hängt vor allem damit zusammen, dass Fahrzeuge in einem vorgeschriebenen zeitlichen Rhythmus zur Hauptuntersuchung vorgeführt werden müssen.

### 3.4.3 Tachostand als Hilfsvariable für eine Fahrleistungsschätzung

Empirische Tachostandsdaten sind letztlich Wertepaare des Typs

$$(z_{it}, k_{it}),$$

wobei  $z_{it}$  den am Fahrzeugprüftag  $(i, t)$  erreichten Tachostand (in km) und  $k_{it}$  das am Fahrzeugprüftag  $(i, t)$  erreichte Alter des Fahrzeugs (in Tagen) bezeichnet. Aus diesen beiden originären Variablen kann bezogen auf das Fahrzeug  $i$  und den Prüftag  $t$  zwar die mittlere bisherige Fahrleistung

$$(3-28) \quad m_{Zit} = z_{it} / k_{it} \quad (\text{km pro Tag})$$

seit dem Tag der Erstanmeldung als neue Variable berechnet werden, nicht aber die aktuelle Fahrleistung  $u_{it}$  (in km) des Fahrzeugs  $i$  in der Bezugsperiode bis zu seinem Prüftag  $t$ , die gemäß Abschnitt 3.3.2 zur Schätzung des periodenbezogenen Fahrleistungstotalwerts  $Y$  benötigt wird. Die eigentlich interessierende Variable  $u_{it}$  ist im Fall der HU-Daten, die dem Erhebungsdesign „Tachostandsablesung an einem Stichtag“ entsprechen, also eine latente, d.h. nicht beobachtbare Variable.

Tachostandsdaten aus Hauptuntersuchungen bieten aber immerhin die Möglichkeit, auf der Basis gewisser Modellannahmen einen Ersatz- oder Näherungswert  $\hat{u}_{it}$  für die nicht beobachtete aktuelle Fahrleistung  $u_{it}$  (Fahrleistung in der Bezugsperiode bis zum Prüftag  $t$ ) zu berechnen. Ein solches Modell hat ganz allgemein die funktionale Form

$$(3-29) \quad \hat{u}_{it} = \varphi(z_{it}, k_{it}),$$

wobei die Transformationsfunktion  $\varphi$  nicht für alle Fahrzeugarten dieselbe sein muss.

Beispiel:

In Abschnitt 3.3.2 wurde gezeigt, dass für ein Fahrzeug  $i$  und einen Stichtag  $t$  innerhalb einer festgelegten Bezugsperiode (z.B. Kalenderjahr 2009) gilt

$$m_{Zit} = m_{Fi} \cdot (d_i / k_{it}) + m_{Uit} \cdot (t / k_{it}).$$

Die mittlere bisherige Fahrleistung  $m_{Zit}$  (mittlere bisherige Fahrleistung des Fahrzeugs  $i$  bis zum Stichtag  $t$  innerhalb der betrachteten Periode) ist gemäß dieser Darstellung gleich dem gewogenen arithmetischen Mittel der mittleren früheren Fahrleistung  $m_{Fi}$  (mittlere Fahrleistung vor der betrach-

teten Periode) und der mittleren aktuellen Fahrleistung  $m_{Uit}$  (mittlere Fahrleistung in der betrachteten Periode bis zum Stichtag  $t$ ), wobei die entsprechenden Lebenszeitanteile  $d_i/k_{it}$  bzw.  $t/k_{it}$  als Gewichtungsfaktoren dienen. Unter der Hypothese (vgl. hierzu auch Abschnitt 6.4)

$$(H) \quad m_{Fi} = m_{Uit},$$

d.h. Gleichheit der mittleren früheren und mittleren aktuellen Fahrleistung, kann mittels der aus den originären Tachostandsdaten gemäß (3-28) direkt abgeleiteten Variablen  $m_{Zit} = z_{it}/k_{it}$  ein Näherungswert für die nicht beobachtete Variable  $u_{it}$  berechnet werden:

$$(3-30) \quad \hat{u}_{it} = \varphi(z_{it}, k_{it}) = t \cdot z_{it}/k_{it} = t \cdot m_{Zit}$$

Unter der Hypothese (H) wird also die "aktuelle Fahrleistung", d.h. die Fahrleistung des Fahrzeugs in der Bezugsperiode bis zum Fahrzeugprüftag gleichgesetzt mit dem Produkt aus mittlerer bisheriger Fahrleistung  $m_{Zit}$  des geprüften Fahrzeugs (in km/Tag) und Länge  $t$  des „aktuellen Zeitraums“ (in Tagen).

Bei der Schätzung des periodenbezogenen Fahrleistungstotalwerts  $Y$  auf der Basis von empirischen Tachostandsdaten ( $z_{it}$ ,  $k_{it}$ ) muss zunächst für jedes Fahrzeug  $i$  in der Stichprobe nach der Formel

$$(3-31) \quad \hat{y}_i = \hat{u}_{iT} = T \cdot z_{it}/k_{it} \quad (i=1, \dots, n)$$

ein Schätzwert für die Fahrleistung  $y_i = u_{iT}$  in der gesamten Bezugsperiode bestimmt werden. Den letztlich interessierenden Schätzwert  $\hat{Y}$  für den periodenbezogenen Fahrleistungstotalwert  $Y$  erhält man dann durch Hochrechnung der „künstlichen“ Stichprobendaten  $\hat{y}_i$  zur Fahrleistung der geprüften Fahrzeuge in der betrachteten Bezugsperiode auf die Gesamtheit aller Fahrzeuge des untersuchten Kollektivs.

Die Validität eines statistischen Verfahrens, mit dessen Hilfe stichtagsbezogene Tachostandsdaten genutzt werden, um periodenbezogene Fahrleistungstotalwerte zu schätzen, hängt natürlich wesentlich davon ab, wie realitätsnah die getroffene Modellannahme (Hypothese) ist. Im obigen Beispiel läuft dies auf die Frage hinaus, wie gut die mittlere bisherige Fahrleistung von Fahrzeugen mit der mittleren aktuellen Fahrleistung von Fahrzeugen übereinstimmt. Die notwendige empirische Prüfung kann natürlich nicht auf der Grundlage von reinen Tachostandsdaten erfolgen. Hierfür sind vielmehr weitere externe Datenbestände erforderlich, bei denen für die untersuchten Fahrzeuge sowohl der aktuelle Tachostand als auch die Fahrleistung in einem aktuellen Zeitraum innerhalb der interessierenden Bezugsperiode erfasst worden ist.

## **4 Prozessdaten aus Kfz-Hauptuntersuchungen als Grundlage einer Fahrleistungsstatistik**

### **4.1 Prozessdaten als empirische Basis statistischer Berichtssysteme**

Prozessdaten, d.h. Daten, die im Rahmen technischer oder wirtschaftlicher Prozesse ohnehin anfallen, können unter bestimmten Bedingungen auch zur Erstellung von Statistiken genutzt werden. Aus Sicht der Statistikproduzenten handelt es sich bei Prozessdaten um Sekundärdaten; im Unterschied zu vielen anderen Arten von Sekundärdaten liegen Prozessdaten aber typischerweise nicht in aggregierter Form sondern als Individualdaten (Mikrodaten) vor. Besonders unter forschungsökonomischen Gesichtspunkten stellen Prozessdaten eine attraktive Alternative zur oft kostspieligen Sammlung von Primärdaten dar.

Im betriebswirtschaftlichen Bereich und speziell in der Wirtschaftsinformatik beschäftigt man sich unter dem Schlagwort „Business Intelligence“ (Intelligence im Sinne von Aufklärung oder Informationsbereitstellung) seit den 1990er Jahren intensiv mit Verfahren und Prozessen zur systematischen Sammlung, Auswertung und Darstellung von entscheidungsrelevanten Daten in elektronischer Form. Die in den ERP-Systemen (Anwendungssoftware zur Unterstützung der Ressourcenplanung eines gesamten Unternehmens) anfallenden Daten werden genutzt, um unter verschiedenen Blickwinkeln die Situation des Unternehmens darzustellen, zu analysieren und ggf. zu bewerten. Die Analyse erfolgt nicht in den ERP-Systemen selbst, sondern in einer davon getrennten Datenbasis, dem Data-Warehouse und ist mit einer Automatisierung des Berichtswesens verbunden.

Auch die statistische Berichterstattung über Strukturen und Entwicklungen in den verschiedenen volkswirtschaftlichen Sektoren stützt sich zunehmend auf Prozessdaten. Eine Vorreiterrolle kann hier der Gesundheits- und Arbeitsmarktstatistik zugeschrieben werden. Aber auch im Verkehrsbereich werden die Bemühungen verstärkt, die Berichtspflichtigen durch Nutzung von Prozessdaten zu entlasten.

Entsprechende Forderungen wurden im Verkehrsbereich im Übrigen schon vor mehr als 25 Jahren formuliert. So finden sich im Tagungsband zum DVWG-Workshop über Verkehrsstatistik von 1984 folgende Feststellungen:

„Zu denken ist in Zukunft insbesondere an: ... - Verstärkter Rückgriff auf Daten aus Verwaltungsunterlagen bzw. auf bereits automatisierte Unterlagen im Bereich der Auskunftspflichtigen. - Verstärkte Anwendung von Schätzmethoden, wenn sich auf diese Weise Primärer-

hebungen vermeiden lassen bzw. Verknüpfung von Daten aus verschiedenen Statistiken durch geeignete Verfahren.“ (Törkel, B., 1985, S. 83).

Die vorliegende Untersuchung setzt die Bemühungen in der oben angesprochenen Richtung fort. Vor der Behandlung des Hochrechnungsverfahrens wird zunächst die Datengrundlage einer möglichen neuen Fahrleistungsstatistik dargestellt.

#### **4.2 Technische Überwachungsinstitutionen als Lieferanten von Prozessdaten für Zwecke einer Fahrleistungsstatistik**

Die Gesamtheit der während eines bestimmten Zeitraums durchgeführten Hauptuntersuchungen (HU) von Kraftfahrzeugen ist ihrer statistischen Natur nach eine Ereignisgesamtheit. Jedes Element dieser Gesamtheit stellt ein „Prüfereignis“ bzw. eine „Prüfung“ dar. Das Ereignis „Prüfung“ besteht in der Vorführung und Untersuchung eines Kfz bei einer Technischen Überwachungsinstitution, im Folgenden auch kurz Prüforganisation genannt. Für jede durchgeführte Kfz-Prüfung wird von der jeweiligen Überwachungsinstitution ein Prüfbericht erstellt, der als Kernmerkmale die Ergebnisse der technischen Prüfung sowie ergänzende Merkmale des Fahrzeugs und seiner Prüfung enthält. Einige dieser ergänzenden Merkmale aus dem Prüfbericht stellen die empirische Basis der geplanten neuen Fahrleistungsstatistik dar.

Um die organisatorischen und datentechnischen Voraussetzungen für das vorliegende Methodenentwicklungsprojekt und die spätere Implementierung der neuen Fahrleistungsstatistik zu schaffen, initiierte das BMVBS eine Vereinbarung zwischen dem Ministerium und einem Kreis von Technischen Überwachungsinstitutionen, in welcher sich diese zur Lieferung von Daten aus Hauptuntersuchungen bereit erklären und in der auch die Modalitäten der Datenbereitstellung für Zwecke einer Fahrleistungsstatistik geregelt werden. Die im Rahmen der genannten Übereinkunft verfügbar gemachten Daten der am Verfahren der Datenlieferung teilnehmenden Prüforganisationen bilden die Rohdaten für die geplante Fahrleistungsstatistik. Im Folgenden wird hier kurz von „HU-Rohdaten“ gesprochen.

Für Zwecke der Entwicklung eines Verfahrens zur Fahrleistungsschätzung wurden in der Anfangsphase des vorliegenden Projekts von 9 Technischen Überwachungsinstitutionen zunächst Rohdaten aus Kfz-Hauptuntersuchungen im Prüfzeitraum 2006 zur Verfügung gestellt. Mit einer Ausnahme lieferten diese Institutionen im weiteren Projektverlauf auch Daten für die Prüfjahre 2007 und 2008. Gegen Projektende kamen schließlich noch HU-Rohdaten einer weiteren Prüforganisation (TÜV Süd) für die drei genannten Jahre hinzu. Bei den insgesamt 10 datenliefernden Prüforganisationen handelt es sich im Einzelnen um folgende Institutionen:



1	Hessen TP
2	Hessen ÜO
3	TÜV Thüringen [nur 2006]
4	TÜV Pfalz UO
5	TÜV Pfalz TP
6	FSP
7	TÜV Rheinland TP
8	TÜV Rheinland UO
9	GTÜ
10	TÜV Süd

Die unter Ziffer 4 bis 8 genannten Prüforganisationen sind dem TÜV Rheinland zugeordnet. Das aus den einzelnen Lieferungen resultierende Datenvolumen liegt pro Prüfjahr in der Größenordnung von ca. 9 Mio. Datensätzen, wobei jeder Datensatz („Fall“) einer im betreffenden Jahr von der betreffenden Prüforganisation durchgeführten Hauptuntersuchung eines Kraftfahrzeugs - im statistischen Sinn also einem Prüfereignis bzw. einer Prüfung - entspricht. Physisch liegen die von den Technischen Überwachungsinstitutionen übermittelten HU-Datenbestände immer für jeweils 1 Halbjahr vor.

Da nicht alle Prüforganisationen am Verfahren der Datenlieferung teilnehmen, handelt es sich bei den hier tatsächlich verfügbaren HU-Rohdaten um eine Stichprobe aus der Menge aller HU-Prüfdaten, die in der Bundesrepublik Deutschland als Prozessdaten anfallen.

### **4.3 Fahrzeug- und prüfterminbezogene Merkmale des HU-Rohdatensatzes**

Entsprechend der getroffenen Vereinbarung enthalten die von den teilnehmenden Prüforganisationen bereitgestellten Rohdaten aus Kfz-Hauptuntersuchungen pro Fall in der Regel die folgenden sechs fahrzeug- bzw. prüfterminbezogenen Merkmale:

- Fahrzeug-Herstellernummer 4-stellig, kurz Herstellercode (Her)
- KBA-Typ-Schlüssel 3-stellig, kurz Typcode (Typ)
- Jahr der Erstanmeldung (Jahr0)
- Monat der Erstanmeldung (Monat0)
- Monat der Hauptuntersuchung (Monat)
- Kilometerstand am Tag der Hauptuntersuchung (Tacho)

Im Vorgriff auf die später im Detail dargestellte Vorgehensweise bei der Zusammenführung und Aufbereitung der einzelnen Rohdatendateien für die Zwecke des vorliegenden Forschungs- und Entwicklungsprojekts sei an dieser Stelle angemerkt, dass jedem Datensatz noch zwei weitere Variablen hinzugefügt werden müssen<sup>13</sup>, um den HU-Rohdatenbestand zu vervollständigen:

- Jahr der Hauptuntersuchung (Jahr)
- Prüforganisation (HUorg)

Der Kilometerstand des Fahrzeugs am Tag der Hauptuntersuchung, der letztlich der Gesamtfahrleistung des Fahrzeugs seit seiner Erstanmeldung (bisherige Gesamtfahrleistung) entspricht, ist natürlich das Kernmerkmal des HU-Datenbestands. In Verbindung mit Jahr und Monat der Hauptuntersuchung sowie Jahr und Monat der Erstanmeldung kann aus dem Kilometerstand am Untersuchungstag nämlich die mittlere bisherige Fahrleistung (in km pro Monat), berechnet werden, dasjenige Merkmal also, welches gemäß Kapitel 3 bei der Fahrleistungsschätzung eine zentrale Rolle spielt.

Aus dem obigen Merkmalskatalog sind aber auch die fahrzeugbezogenen Variablen Herstellercode (Her) und Typcode (Typ) unverzichtbar, da mit ihrer Hilfe unter anderem die beiden neuen Variablen

- Antriebsart und
- Fahrzeugart

gebildet werden können. Diesen neuen Variablen kommt, wie noch gezeigt wird, wegen der bei der Hochrechnung notwendigen Verknüpfung von HU-Daten mit ZFZR-Bestandsdaten, aber auch im Hinblick auf das Tabellenprogramm einer auf HU-Daten basierenden Fahrleistungsstatistik eine Schlüsselrolle zu.

Als Testdaten für die Methodenentwicklung aber auch als Eingangsmaterial einer späteren laufenden Fahrleistungsstatistik müssen die von den teilnehmenden Prüforganisationen bereitgestellten HU-Rohdaten in geeigneter Weise geprüft, plausibilisiert und aufbereitet werden. Das hierfür entwickelte Verfahren und die Ergebnisse der Anwendung des Verfahrens auf die HU-Daten der Prüfjahre 2006 und 2007 werden in Kapitel 8 dargestellt.

---

<sup>13</sup> Teilweise ist im Datensatz auch noch der Zulassungsbezirk des Fahrzeugs als Merkmal vorhanden.

## **5 Methodische Grundsatzfragen der Eignung von HU-Daten als Basis einer Fahrleistungsstatistik**

### **5.1 Die methodischen Probleme im Überblick**

Mit der Zielsetzung des vorliegenden Projekts, ein statistisches Verfahren zur Hochrechnung kalenderjahrspezifischer Fahrleistungen für die in der Bundesrepublik Deutschland mit amtlichem Kennzeichen zugelassenen Kraftfahrzeuge zu entwickeln, ist die Grundgesamtheit, für die der Jahrestotalwert der Fahrleistungen zu schätzen ist, durch die Gesamtheit der im ZFZR registrierten Kfz gegeben. Die Stichprobe, auf deren Grundlage dies geschehen soll, ist durch die von den datenliefernden Prüforganisationen für aufeinander folgende Prüffahre bereitgestellten HU-Rohdaten gegeben, welche für die zur Hauptuntersuchung vorgeführten Fahrzeuge den Kilometerstand zum Prüftermin als Kernmerkmal enthalten.

Vor diesem Hintergrund stellen sich verschiedene methodische Grundsatzfragen:

- Inwieweit ist es überhaupt möglich, einen kalenderjahrspezifischen Fahrleistungstotalwert auf der Grundlage einer Stichprobe zu schätzen, die gar keine direkte Fahrleistungsinformationen für das betreffende Kalenderjahr enthält sondern nur stichtagsbezogene Informationen zum erreichten Tachostand?
- Kann die Veränderung des Fahrleistungstotalwertes von einem Kalenderjahr zum nächsten auf der Basis stichtagsbezogener Tachostandsdaten – ggf. aus mehreren aufeinander folgenden Prüffahren – geschätzt werden?
- Können Tachostands- und daraus abgeleitete Fahrleistungsinformationen, die sich auf die in einem bestimmten Kalenderjahr geprüften (d.h. zur Hauptuntersuchung vorgeführten) Fahrzeuge beziehen, auf die im jeweiligen Jahr nicht geprüften Fahrzeuge übertragen werden?
- Sind die Fahrzeuge, die bei den datenliefernden Prüforganisationen zur Hauptuntersuchung vorgeführt werden, repräsentativ für alle im jeweiligen Jahr in der Bundesrepublik Deutschland geprüften Fahrzeuge?

Im vorliegenden Kapitel wird zunächst die Frage der Repräsentativität der bereitgestellten HU-Daten für alle in der Bundesrepublik Deutschland durchgeführten Kfz-Hauptuntersuchungen behandelt. Im Anschluss daran wird untersucht, welche Probleme daraus resultieren, dass in jeder Bezugsperiode der Fahrleistungsschätzung Tachostandsdaten immer nur für diejenigen Fahrzeuge vorliegen, die in der betreffenden Periode zu einer Hauptuntersuchung vorgeführt worden sind. Den methodischen Ansätzen zur Schätzung periodenbezogener Fahrleistungen sind die Kapitel 6 und 7 gewidmet.

## 5.2 Repräsentativität der bei den datenliefernden Prüforganisationen untersuchten Fahrzeuge für alle untersuchten Fahrzeuge

Gemäß der sog. FU-Statistik des KBA („Fahrzeuguntersuchungen, Hauptuntersuchungen und Einzelabnahmen“) wurden im Jahr 2007 rund 22,3 Mio. insgesamt Hauptuntersuchungen durchgeführt. Für das Jahr 2006 wird diese Zahl mit 21,9 Mio. angegeben. Die zur Entwicklung und Erprobung eines Hochrechnungsverfahrens verfügbaren HU-Rohdatenbestände der in Kapitel 4 genannten datenliefernden Prüforganisationen stellen mit ihren jährlich ca. 9 Mio. Datensätzen demgemäß eine Stichprobe mit einem Auswahlsatz von rund 40 % aus der Gesamtheit aller im betreffenden Jahr untersuchten Fahrzeuge dar.

### 5.2.1 Strukturvergleich zwischen HU-Stichprobe und FU-Statistik

Ob die verwendete HU-Stichprobe ( $n \approx 9$  Mio. Fahrzeuge) für die entsprechende HU-Grundgesamtheit ( $N \approx 22$  Mio. Fahrzeuge) repräsentativ ist, kann im Hinblick auf das Merkmal „Fahrzeugalter zum Zeitpunkt der Prüfung“ durch einen entsprechenden Strukturvergleich geprüft werden. Unterschieden wird in der FU-Statistik nach insgesamt vier, gegenüber der hier gewählten Kategorisierung teilweise aggregierten, Fahrzeugarten: „Kraftfahrzeuge“ (Kraftfahrzeuge), „Pkw“ (Pkw), „Kraftomnibusse, Lkw und sonstige Kfz“, (Kraftomnibusse, Lkw, Sonstige Kfz) sowie „Zugmaschinen“ (Sattelzugmaschinen, sonstige Zugmaschinen, Zugmaschinen LoF). Die folgende Tabelle 1 gibt einen Überblick über die Struktur der Grundgesamtheit der Kfz-Hauptuntersuchungen der Jahre 2006 und 2007.

**Tabelle 1: Grundgesamtheit aller Fahrzeuge bei Hauptuntersuchungen 2006 und 2007 gegliedert nach Fahrzeugart und Fahrzeugalter (FU-Statistik)**

Alter in Jahren	Pkw 2007	Pkw 2006	Kraftfahrzeuge 2007	Kraftfahrzeuge 2006	Lkw, Bus, Sonstige Kfz 2007	Lkw, Bus, Sonstige Kfz 2006	Zugmaschinen 2007	Zugmaschinen 2006
bis 3	2878857	2943424	167762	180318	364258	347639	116637	104883
4 bis 5	2746414	2826076	169580	174356	226077	235032	54828	54317
6 bis 7	2695564	3040779	177418	196748	224937	237258	45361	46499
8 bis 9	2808861	2773110	200601	211482	200342	184955	39490	37251
10+	6904320	6273432	846233	726050	638955	624093	696577	660033
Gesamt	18034016	17856821	1561594	1488954	1654569	1628977	952893	902983

Um die Repräsentativität der vorliegenden HU-Rohdaten für die Gesamtheit aller durchgeführten Kfz-Hauptuntersuchungen hinsichtlich des Fahrzeugalters beurteilen zu können, wurde die HU-Stichprobe nach dem Merkmal „Prüfalter in Jahren“ gemäß Abschnitt 8.5.1 aufgegliedert, welches zu der in der FU-Statistik verwendeten Definition des Fahrzeugalters kompatibel ist. Tabelle 2 zeigt

beispielhaft für die Fahrzeugart Pkw und das Untersuchungsjahr 2007 die Ergebnisse des Strukturvergleichs.

**Tabelle 2: Grundgesamtheit 2007 und Stichprobe 2007 der Pkw bei Hauptuntersuchungen gegliedert nach Fahrzeugalter**

Alter	Grundgesamtheit (FU-Statistik)	HU-Stichprobe	Grundgesamtheit in %	HU-Stichprobe in %	Differenz in Prozentpunkten
bis 3 Jahre	2.878.857	1.307.092	16,0%	17,5%	1,5%
4 bis 5 Jahre	2.746.414	1.117.143	15,2%	14,9%	-0,3%
6 bis 7 Jahre	2.695.564	1.134.272	14,9%	15,2%	0,2%
8 bis 9 Jahre	2.808.861	1.184.707	15,6%	15,8%	0,2%
10+ Jahre	6.904.320	2.743.618	38,3%	36,6%	-1,6%
Gesamt	18.034.016	7.486.832	100,0%	100,0%	0,0%

Im Ergebnis zeigt sich zwischen der Grundgesamtheit aller Kfz-Hauptuntersuchungen des Jahres 2007 und der hier vorliegenden Stichprobe von Kfz-Hauptuntersuchungen 2007 eine gute Übereinstimmung, so dass zumindest im Hinblick auf die Altersstruktur der geprüften Fahrzeuge davon ausgegangen werden kann, dass die HU-Rohdaten der datenliefernden Prüforganisationen alle in der Bundesrepublik Deutschland durchgeführten Kfz-Hauptuntersuchungen eines Kalenderjahres repräsentieren.

### 5.2.2 Validität der HU-Daten bezüglich der Untersuchungsintervalle

Aus einem anderen Blickwinkel kann die Validität der hier genutzten HU-Stichprobe 2007 in gewissem Umfang auch dadurch beurteilt werden, dass differenziert nach dem Jahr der Erstzulassung das Durchschnittsalter der im Jahr 2007 bei den datenliefernden Prüforganisationen zur Hauptuntersuchung vorgeführten Fahrzeuge berechnet wird (Alter zum Prüfzeitpunkt in Monaten).

Tabelle 3 zeigt deutlich, dass im Hinblick auf das Durchschnittsalter zum Prüfzeitpunkt ab dem Erstzulassungsjahr 2006 gut die Vielfachen von 12 Monaten erreicht werden, wie dies wegen der bestehenden StVZO-Bestimmungen zu den Untersuchungsintervallen zu erwarten ist. So sind beispielsweise in der Fahrzeugart „Pkw“ die im Jahr 2007 geprüften Kfz mit dem Erstanmeldejahr 2004 („Pkw-Kohorte 2004“) bei der Prüfung im Durchschnitt 36 Monate alt. In wenigen Fällen weicht das dargestellte Durchschnittsalter in Monaten von den Erwartungen ab (z.B. 25 Monate für die Pkw-Kohorte 2005), dies wird hier aber durch die Rundung der Mittelwerte verursacht, der Wert mit zwei Nachkommastellen beträgt in diesem Fall 24,60 Monate.

**Tabelle 3: Durchschnittsalter der in der HU-Stichprobe erfassten Fahrzeuge bei der Hauptuntersuchung 2007 gegliedert nach Erstzulassungsjahr**

Zulassungsjahr	Durchschnittsalter in Monaten bei der Hauptuntersuchung 2007				
	Lkw	Sattelzugmaschinen	Omnibusse	Krafträder	Pkw
2007	5	7	7	4	6
2006	13	12	12	13	12
2005	24	24	24	25	25
2004	35	36	36	35	36
2003	48	48	47	49	47
2002	59	60	60	59	60
2001	72	72	72	73	71
2000	83	84	83	83	84
1999	96	96	96	97	96
1998	107	108	108	108	108
1997	120	119	119	121	120
1996	131		132	131	132
1995	144		144	145	143
1994	154		156	156	157
1993	168		168	169	167
1992	179		180	180	181
1991	192		192	194	191
1990	203		204	205	205
1989	216		217	218	214
1988	227		228	229	229
1987	240		240	241	238

Vor diesem Hintergrund kann man davon ausgehen, dass die in der HU-Stichprobe erfassten Kraftfahrzeuge (genauer: Kfz-Hauptuntersuchungen) im Hinblick auf die Periodizität der Vorführung zur Hauptuntersuchung repräsentativ für die Gesamtheit aller geprüften Kraftfahrzeuge sind und dies wohl auch in Zukunft so sein wird – zumindest solange keine datenliefernden Institutionen wegfallen oder hinzukommen.

## 5.3 Repräsentativität der in einer Periode geprüften Fahrzeuge für den betreffenden Fahrzeugbestand insgesamt

### 5.3.1 Beschränkung auf geprüfte Fahrzeuge als Coverage-Problem

Obwohl die für eine Fahrleistungsstatistik bereitgestellten HU-Daten naturgemäß eine Stichprobe aus der Grundgesamtheit der *geprüften* Fahrzeuge darstellen, soll diese Stichprobe auf die Gesamtheit *aller* Fahrzeuge, einschließlich der im betreffenden Jahr *nicht geprüften* Fahrzeuge, hochgerechnet werden<sup>14</sup>. Die Stichprobe, um deren Hochrechnung es geht, ist also keine Stichprobe aus der Grundgesamtheit, auf welche hochgerechnet werden soll. Fahrzeuge, bei denen im Bezugsjahr keine Hauptuntersuchung durchgeführt wird, haben im Hinblick auf die HU-Stichprobe die Auswahlwahrscheinlichkeit Null.

Aus methodischer Sicht liegt hier ein sog. „Coverage-Problem“ vor, da die „Auswahlgrundlage“, aus welcher die HU-Stichprobe „gezogen“ wird (im Prinzip ist dies die FU-Statistik), einen Teil der interessierenden Grundgesamtheit, nämlich die im jeweiligen Jahr *nicht* geprüften Fahrzeuge, nicht überdeckt. Wie groß der nicht überdeckte Teil der eigentlichen „Zielgesamtheit“ ist, kann wegen der Nichterfassung der vorübergehend abgemeldeten Fahrzeuge im ZFZR im Übrigen nur näherungsweise angegeben werden. Die Zahl der am Stichtag 1.1.2008 angemeldeten Fahrzeuge liegt bei rund 49,3 Mio. Kfz und die Zahl der im Verlauf des Jahres 2007 durchgeführten Hauptuntersuchungen bei 22,2 Mio.. Damit hatten im Jahr 2007 mehr als die Hälfte aller Fahrzeuge keine Chance, in die HU-Stichprobe zu kommen.

Die Vorführung oder Nichtvorführung eines Fahrzeugs zur Hauptuntersuchung in einer bestimmten Periode kann durchaus in einem statistischen Zusammenhang mit dem Tachostand stehen. Betrachtet man eine bestimmte Fahrzeugkohorte, etwa alle Pkw mit dem Erstzulassungsjahr 1999, so wird entsprechend der StVZO-Untersuchungsintervalle ein Großteil dieser Fahrzeuge im Jahr 2002 erstmals bei einer Hauptuntersuchung erscheinen und danach wieder im zweijährigen Turnus (gemeint sind hier natürlich nur die jeweils „Überlebenden“). Bis zum Ende des Jahres 2009 werden die meisten (der bis dahin überlebenden) Fahrzeuge aus der Pkw-Kohorte 1999 demnach viermal (nämlich 2002, 2004, 2006 und 2008) zu einer Hauptuntersuchung vorgeführt worden sein. Es wird in der Kohorte aber auch Fahrzeuge geben, die im fraglichen Zeitraum häufiger, etwa 5-, 6- oder gar 7-mal, bei einer Prüforganisation vorgeführt worden sind.

---

<sup>14</sup> Sieht man von der Möglichkeit ab, dass ein Kfz im Laufe eines Kalenderjahres mehrmals zu einer Hauptuntersuchung vorgeführt wird (was nur in sehr begrenzten Ausnahmefällen vorkommen dürfte), so kann eine Hauptuntersuchung mit einem (geprüften) Kfz gleichgesetzt werden.

Fahrzeuge mit höherer Prüffrequenz haben natürlich eine überdurchschnittlich hohe Wahrscheinlichkeit, in der HU-Stichprobe zu erscheinen (sog. Inklusionswahrscheinlichkeit). Wenn Fahrzeuge mit überdurchschnittlicher Inklusionswahrscheinlichkeit einen vom Durchschnitt der Kohorte abweichenden (z.B. höheren) Tachostand aufweisen<sup>15</sup>, so ist der auf Basis der HU-Stichprobe eines beliebigen Bezugsjahres aus dem Zeitraum 1999 bis 2009 ermittelte durchschnittliche Tachostand – und damit auch die daraus abgeleitete periodenspezifische Fahrleistung – verzerrt. Diese Verzerrung (z.B. nach oben) der für die Jahre 1999 bis 2009 ermittelten durchschnittlichen Tachostände bzw. Fahrleistungen betrifft natürlich nicht nur die Pkw-Kohorte 1999, sondern alle Pkw-Kohorten, deren „Lebenszeit“ den Kalenderzeitraum 1999 bis 2009 ganz oder teilweise überdeckt.

### 5.3.2 Strategien zur Reduzierung von Verzerrungen

Dem angesprochenen Coverage-Problem muss bei der Fahrleistungsschätzung, soweit irgendwie möglich, Rechnung getragen werden. Da die hierfür eigentlich notwendigen empirischen Daten – nämlich Tachostand bzw. Fahrleistung in Abhängigkeit von der Prüffrequenz des Fahrzeugs – nicht vorliegen, kann dies nur in grober Näherung erfolgen. Wie später noch gezeigt wird, geschieht dies beim letztlich gewählten Hochrechnungsverfahren dadurch, dass

- im Hochrechnungsrahmen für jedes Bezugsjahr der Fahrleistungsschätzung der Bestand der Fahrzeuge mit und ohne Hauptuntersuchung im betreffenden Jahr gesondert erscheint und
- bei der Hochrechnung der HU-Stichprobe zwischen „regelmäßig (regulär)“ und „nicht regelmäßig (irregulär)“ vorgeführten Fahrzeugen unterschieden wird<sup>16</sup>, da für nicht regelmäßig vorgeführte Fahrzeuge eine tendenziell höhere Prüffrequenz unterstellt werden kann (und im Übrigen auch eine überdurchschnittlich hohe Fahrleistung nachweisbar ist).

---

<sup>15</sup> bei gegebenem Stichtag der Tachostandsablesung

<sup>16</sup> Gemeint ist hier, ob das Fahrzeug im „normalen Rhythmus“ oder „außerhalb des normalen Rhythmus“ zur HU vorgeführt wird. Fahrzeuge aus der Pkw-Kohorte 1999, die in den Jahren 2002, 2004, 2006 und 2008 zu einer Hauptuntersuchung vorgeführt worden sind, gelten in diesen Jahren als „regelmäßig“ vorgeführte Fahrzeuge. Wird ein Fahrzeug aus der Pkw-Kohorte 1999 z.B. im Jahr 2000 oder im Jahr 2007 einer Hauptuntersuchung unterzogen, so handelt es sich bei dieser Hauptuntersuchung um eine „nicht regelmäßige“ Prüfung.



## 6 Alternative Modellansätze für die Schätzung periodenbezogener Fahrleistungen aus Tachostandsdaten

Die zentrale Frage bei der Hochrechnung des Fahrleistungstotal für eine aktuelle Bezugsperiode (Kalenderjahr) betrifft die Modellannahmen, unter denen die „aktuelle“ Jahresfahrleistung aus Daten zur gesamten bisherigen Fahrleistung von Fahrzeugen berechnet wird. Grundsätzlich lassen sich hier verschiedene mögliche Modellansätze unterscheiden:

M1: Schätzung der aktuellen Fahrleistung der Fahrzeuge einer bestimmten Altersgruppe in der Bezugsperiode unter Verwendung der auf einen Stichtag aus dieser Periode bezogenen bisherigen Gesamtfahrleistung „benachbarter“ Altersgruppen (kohortenübergreifende Schätzung auf der Basis von Querschnittsdaten)

M2: Schätzung der aktuellen Fahrleistung der Fahrzeuge einer bestimmten Altersgruppe in der Bezugsperiode unter Verwendung eines Regressionsmodells für den Tachostand als Funktion des Fahrzeugalters (regressionsanalytische Schätzung auf der Basis von Querschnittsdaten)

M3: Schätzung der aktuellen Fahrleistung der Fahrzeuge einer bestimmten Altersgruppe in der Bezugsperiode unter Verwendung der bisherigen Gesamtfahrleistung der betreffenden Fahrzeugkohorte in aufeinanderfolgenden Perioden (kohortenspezifische Schätzung auf der Basis von Längsschnittsdaten).

M4: Schätzung der aktuellen Fahrleistung der Fahrzeuge in der Bezugsperiode unter Verwendung der auf diese Periode bezogenen bisherigen Lebensdauer und bisherigen Gesamtfahrleistung der Fahrzeuge (endogene Schätzung auf der Basis von Querschnittsdaten).

Die Validität und Praktikabilität der genannten Ansätze als methodische Grundlage einer neuen Fahrleistungsstatistik wird im Folgenden untersucht und bewertet.

### 6.1 Kohortenübergreifende Schätzung auf der Basis von Querschnittsdaten

Prinzipiell können kalenderjahrbezogene Fahrleistungen geschätzt werden, indem die Differenz der durchschnittlichen Tachostände zweier aufeinanderfolgender Altersklassen gebildet wird. Es sei  $\mu_z(t, a)$  der auf das Jahr  $t$  bezogene mittlere Tachostand der Fahrzeuge, die im Kalenderjahr  $e=t-a$  erstmals zugelassen wurden und somit im „aktuellen“ Jahr  $t$  gerade  $a$  Jahre alt sind ( $a=0, 1, 2, \dots$ ). Bei einer auf Querschnittsdaten, d.h. Tachostandsdaten aus dem Jahr  $t$  beruhenden „kohortenübergreifenden“ Herangehensweise an die Fahrleistungsschätzung (Modell M1) wird die Differenz

$$(6-1) \quad \mu_z(t, a+1) - \mu_z(t, a)$$

der altersgruppenspezifischen mittleren Tachostände im Jahr  $t$  als Schätzwert für die mittlere Fahrleistung  $\mu_y(t, a)$  der  $a$ -jährigen Fahrzeuge im Jahr  $t$  verwendet (Fahrleistung in km/Jahr).

Durch Multiplikation der obigen Mittelwertdifferenz mit dem mittleren Bestand  $\mu_B(t, a)$  der  $a$ -jährigen Fahrzeuge im Jahr  $t$  erhält man mit

$$(6-2) \quad [\mu_z(t, a+1) - \mu_z(t, a)] \cdot \mu_B(t, a)$$

eine Schätzung für die Gesamtfahrleistung  $Y(t, a)$  der  $a$ -jährigen Fahrzeuge im Jahr  $t$ .

Diese Schätzung basiert ausschließlich auf Daten aus der Bezugsperiode  $t$  (also Querschnittsdaten) und beruht in gewisser Weise auf der Annahme, dass der von den Fahrzeugen aus dem Erstzulassungsjahr  $t-a-1$  im Bezugsjahr  $t$  bereits erreichte mittlere Tachostand annähernd genauso groß ist wie der mittlere Tachostand, den die Fahrzeuge aus dem (späteren) Erstzulassungsjahr  $t-a$  im Jahr  $t+1$  erreicht haben werden (der letztgenannte mittlere Tachostand ist im Jahr  $t$ , dem Bezugsjahr der Fahrleistungsschätzung, natürlich noch nicht bekannt).

Dieser zunächst plausibel erscheinende Ansatz wurde im Rahmen der Methodenentwicklung ausführlich getestet, letztlich aber doch als nicht praktikabel verworfen. Der Grund hierfür liegt darin, dass mit zunehmendem Alter der Fahrzeuge überproportional viele Kfz mit hohen bisherigen Fahrleistungen aus dem Bestand herausfallen<sup>17</sup>. Als Folge hiervon weisen die durchschnittlichen Tachostände der oberen Altersgruppen im Vergleich zu den jeweils darunterliegenden Altersgruppen im Hinblick auf die hier vorliegende Fragestellung „zu geringe“ Werte auf, womit die durch (6-1) gegebene Schätzung

$$\mu_z(t, a+1) - \mu_z(t, a)$$

für die interessierende „aktuelle“ mittlere Fahrleistung  $\mu_y(t, a)$  tendenziell zu niedrig ausfällt (in manchen Fällen werden die Differenzen der mittleren Tachostände aufeinanderfolgender Altersgruppen bzw. Kohorten sogar negativ).

---

<sup>17</sup> Bei gegebenem Fahrzeualter ist die Wahrscheinlichkeit für das Ausscheiden eines Fahrzeugs aus dem Bestand umso höher, je höher der Tachostand des Fahrzeugs ist.

## 6.2 Regressionsanalytische Schätzung auf der Basis von Querschnittsdaten

Die oben beschriebene Differenzenbildung zur Schätzung der periodenbezogenen Fahrleistung von Kraftfahrzeugen liegt in gewisser Weise auch einem regressionsanalytischen Schätzverfahren zugrunde, bei dem der mittlere Tachostand  $Z$  eines Fahrzeugkollektivs als Funktion  $Z=Z(a)$  des Fahrzeugalters  $a$  beim Prüftermin betrachtet wird (Modell M2). Geschätzt werden die Parameter der aggregierten Tachostandsfunktion  $Z(a)$  mit HU-Daten eines oder mehrerer Prüffahre.

Die 1. Ableitung  $Z'(a)$  der Tachostandsfunktion an einer Stelle  $a=\tilde{a}$  (Kfz-Alter) entspricht der Steigung der Tachostandsfunktion und damit (sofern die Funktion  $Z(a)$  monoton wachsend ist) der Zunahme des mittleren Tachostands bei einer infinitesimalen Zunahme des Fahrzeugalters. Die Zunahme des Tachostands beim Anstieg des Fahrzeugalters von  $a$  auf  $a+1$  Jahre ist näherungsweise gleich  $Z'(a)$ . Unterstellt man, dass die Zunahme des mittleren Tachostands beim Übergang vom Alter  $a$  zum Alter  $a+1$  (Alter in Jahren) gleich der in der Bezugsperiode  $t$  mit der Länge 1 Jahr erbrachten Fahrleistung der im Jahr  $t-a$  erstmals zugelassenen Fahrzeuge entspricht, so hat man mit  $Z'(a)$  eine Schätzung der aktuellen mittleren Jahresfahrleistung der betreffenden Kohorte.

Die beiden folgenden Abbildungen zeigen diesen Ansatz beispielhaft für die beiden Fahrzeugarten „Krafträder“ und „Sattelzugmaschinen“. Als Regressionsfunktionen wurden jeweils Polynome gewählt, die eine sehr gute Anpassung an die empirischen Daten gewährleisten (Abszisse: Fahrzeugalter in Jahren; Ordinate: mittlerer Tachostand (links) bzw. mittlere Jahresfahrleistung (rechts) in km).

**Abbildung 1: Regressionsanalytische Schätzung der Jahresfahrleistung 2007 von Krafträdern**

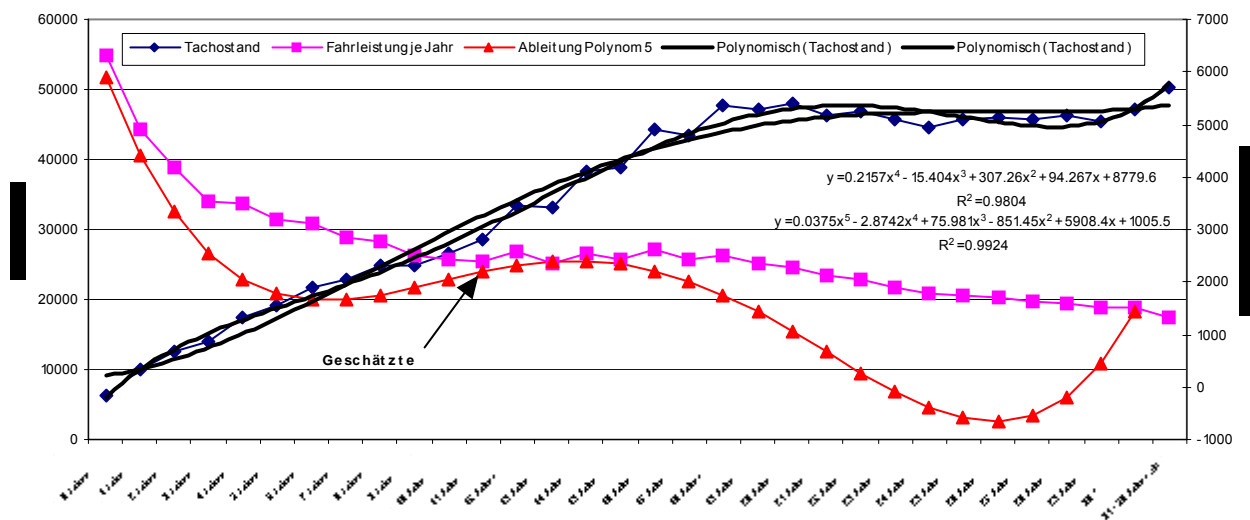
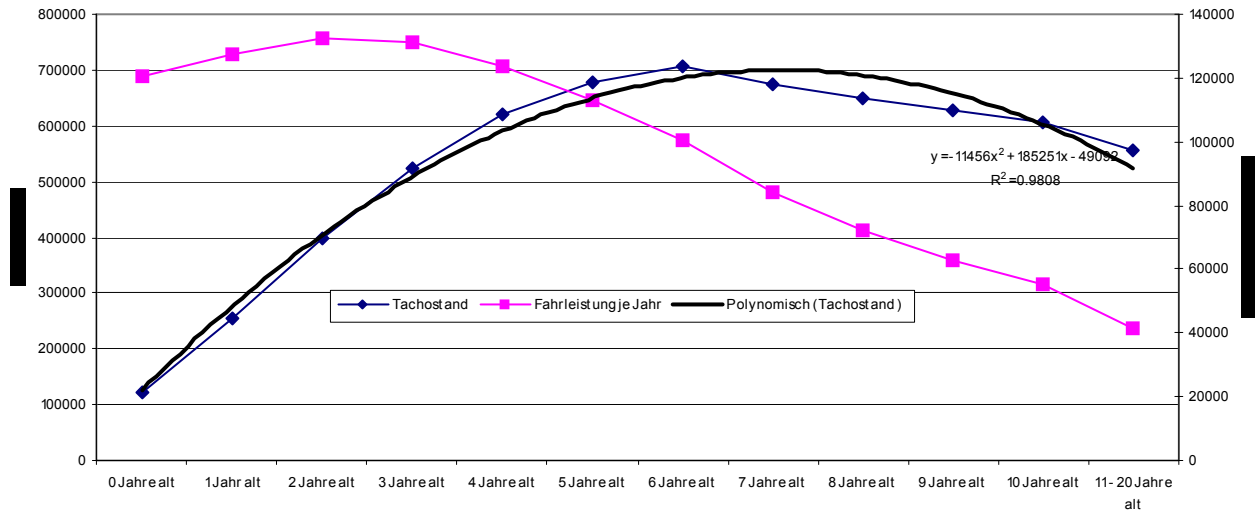


Abbildung 2: Regressionsanalytische Schätzung der Jahresfahrleistung 2007 von Sattelzugmaschinen



Die Abbildungen zeigen die Schwächen des regressionsanalytischen Ansatzes M2 deutlich auf. Wie beim Modell M1 besteht die Grundproblematik darin, dass die aggregierte Tachostandsfunktion  $Z(a)$  nicht monoton wächst, sondern für hohe Werte des Lebensalters  $a$  wieder sinkt. Ganz besonders ausgeprägt ist dieses Phänomen in Abbildung 2 (Sattelzugmaschinen 2007). Fahrleistungsschätzungen mit Hilfe der ersten Ableitung der Tachostandsfunktion können also (vor allem bei älteren Kohorten) zu negativen Jahresfahrleistungen führen, die selbstverständlich nicht plausibel sind.

Ein weiterer Nachteil der Regressionsmethode liegt darin, dass je Fahrzeugart die „beste“ funktionale Verknüpfung zwischen Tachostand und Fahrzeugalter für jedes Prüffahr neu bestimmt werden muss. Eine pauschale Übertragung einer einmal geschätzten Funktion auf alle Prüffahre scheint nicht sinnvoll. Problematisch ist auch, dass es auch keine theoretische Rechtfertigung für das jeweils gewählte Polynom gibt.

### 6.3 Kohortenspezifische Schätzung auf der Basis von Längsschnittdaten

Bei der Verwendung der HU-Daten aufeinanderfolgender Perioden (also Längsschnittdaten<sup>18</sup>) zur Schätzung der Fahrleistung einer bestimmten Fahrzeugkohorte in der Bezugsperiode treten im Wesentlichen dieselben Probleme auf wie bei der kohortenübergreifenden Schätzung auf der Basis von Querschnittsdaten (Modell M1). Dem „periodenübergreifende“ Modellansatz M3 liegt die Annahme zugrunde, dass die nicht erhobene periodenbezogene mittlere Fahrleistung  $\mu_V(t, a)$  durch die Differenz zweier Tachostandsmittelwerte, nämlich

$$(6-3) \quad \mu_Z(t+1, a+1) - \mu_Z(t, a)$$

geschätzt werden kann, wobei die beiden Tachostandsmittelwerte sich auf dieselbe Kohorte (nämlich das Erstzulassungsjahr  $t-a$ ) aber verschiedene Zeitjahre (nämlich die Jahre  $t$  und  $t+1$ ) beziehen.

Auf der Mikroebene, d.h. bei Betrachtung einzelner Fahrzeuge, trifft es natürlich zu, dass der Tachostand im Zeitverlauf nur zunehmen oder konstant bleiben kann<sup>19</sup>. Im Aggregat, d.h. für eine Fahrzeugkohorte, die über die Zeit hinweg beobachtet wird, gilt dies aber nicht: Bedingt durch das tendenziell frühere Ausscheiden von stark genutzten Fahrzeugen aus dem Bestand kommt es beim Vergleich der mittleren Tachostände einer bestimmten Fahrzeugkohorte für aufeinanderfolgende Jahre zu Ergebnissen, die im Hinblick auf die hier vorliegende Aufgabenstellung unplausibel sind (wieder bis hin zu negativen Differenzen). Aus diesem Grund ist es beim gegenwärtigen HU-Merkmalsskatalog<sup>20</sup> nicht Erfolg versprechend, eine Fahrleistungsstatistik auf dem vom Konzept her dynamischen Modellansatz M3 aufzubauen.

Anders wäre natürlich die Situation, wenn eine echte Kohortenanalyse mit Tachostandsdaten möglich wäre. Dies würde aber voraussetzen, dass die HU-Daten in Form von verbundenen Stichproben<sup>21</sup> vorliegen, was tatsächlich nicht der Fall ist.

---

<sup>18</sup> HU-Daten aufeinanderfolgender Jahre sind Längsschnittdaten, jedoch keine Paneldaten. Dies hängt damit zusammen, dass es im verfügbaren HU-Datenbestand kein zeitlich invariantes Identifikationsmerkmal gibt, anhand dessen Ablesungen des Tachostands desselben Fahrzeugs in verschiedenen Jahren zusammengeführt werden könnten.

<sup>19</sup> Von möglichen Manipulationen des Tachostands wird hier einmal abgesehen.

<sup>20</sup> Fehlen eines zeitinvarianten Fahrzeugidentifikationsmerkmals

<sup>21</sup> Eine verbundene Stichprobe läge vor, wenn für ein und dasselbe Fahrzeug die Tachostände bei aufeinanderfolgenden Hauptuntersuchungen registriert und gespeichert würden.

## 6.4 Endogene Schätzung auf der Basis von Querschnittsdaten: Die Durchschnittswertmethode

### 6.4.1 Grundgedanke

Datenbestände aus Hauptuntersuchungen beinhalten, wenn man von weiteren Fahrzeug- und Haltermerkmalen einmal absieht, die folgenden drei Kernmerkmale:

- Jahr und Monat der Erstzulassung
- Jahr und Monat der Hauptuntersuchung
- Kilometerstand zum Zeitpunkt der Hauptuntersuchung.

Die HU-Datenbestände liegen getrennt nach Untersuchungsjahren vor. Dabei versteht es sich, dass sich innerhalb jedes jahresspezifischen Datenbestands Tachostandsangaben zu Fahrzeugen aus unterschiedlichen Kohorten (Erstzulassungsjahre) finden.

Die bisher betrachteten Modellansätze sind alle dadurch gekennzeichnet, dass zur Schätzung der Fahrleistung  $Y(t, a)$  der  $a$ -jährigen Fahrzeuge im Bezugsjahr  $t$  gewisse „exogene“ Tachostandsdaten verwendet werden. Es handelt sich dabei um Daten zum Tachostand, die

- bei gegebenem Bezugsjahr sich auf Fahrzeuge beziehen, die im Bezugsjahr ein von  $a$  verschiedenes Alter haben, d.h. zu einer *anderen Kohorte* gehören (Modelle M1 und M2) bzw. um Daten, die
- bei gegebenem Alter der Fahrzeuge im Bezugsjahr, d.h. bei gegebener Fahrzeugkohorte, den Tachostand der betrachteten Kohorte in *anderen Zeitjahren* betreffen (Modell M3).

Beim jetzt betrachteten Modell M4 wird bei der Schätzung der periodenbezogenen Fahrleistung  $Y(t, a)$  auf solche exogenen Tachostandsdaten (Daten zu anderen Kohorten oder anderen Jahren) komplett verzichtet. Stattdessen wird für die im Bezugsjahr  $t$  gerade  $a$ -jährigen Fahrzeuge das Verhältnis

$$(6-4) \quad \mu_Z(t, a) / a,$$

d.h. die mittlere bisherige Fahrleistung dieser Fahrzeuge, berechnet und als Näherungswert für die eigentlich interessierende mittlere Fahrleistung  $\mu_Y(t, a)$  der  $a$ -jährigen Fahrzeuge im Jahr  $t$  verwendet. Nach Multiplikation mit dem entsprechenden mittleren Fahrzeugbestand hat man mit

$$(6-5) \quad \mu_B(t, a) \cdot \mu_Z(t, a) / a$$

die gesuchte Schätzung für die periodenbezogene Fahrleistung  $Y(t, a)$ . Da bei der Schätzung von  $Y(t, a)$  nur Daten verwendet werden, die sich auf die Kombination  $(t, a)$  beziehen, kann beim Modell M4 von einer „endogenen“ Schätzung auf der Basis von Querschnittsdaten gesprochen werden.

Bei der praktischen Umsetzung der endogenen Schätzung mit Hilfe von HU-Daten kann man auf Einzeldatenebene, d.h. für alle Fahrzeuge  $i$  in der HU-Stichprobe des Prüffjahres  $t$ , das Verhältnis zwischen Tachostand  $z_i$  (in km) und Fahrzeugalter  $k_i$  (in Monaten) berechnet werden. Die Multiplikation der so berechneten mittleren bisherigen Fahrleistung  $z_i/k_i$  (in km/Monat) mit der Länge  $T$  der Bezugsperiode (in Monaten) stellt dann einen Näherungswert  $\hat{y}_i$  für die Fahrleistung  $y_i$  des betreffenden Fahrzeugs in der Bezugsperiode der Hochrechnung dar:

$$(6-6) \quad \hat{y}_i = T \cdot z_i / k_i \quad (i=1, \dots, n)$$

Wegen der zentralen Rolle der mittleren bisherigen Fahrleistung wird das endogene Schätzverfahren für das periodenbezogene Fahrleistungstotal im Folgenden kurz „Durchschnittswertmethode“ genannt. Bezugsperiode ist im Fall der HU-Daten jeweils ein Kalenderjahr, d.h.  $T=12$ .

Anmerkung: Die Durchschnittswertmethode kann aus statistischer Sicht auch als ein Imputationsverfahren, d.h. als Verfahren zur Ergänzung fehlender Werte in einem Datenbestand betrachtet werden: Das eigentlich benötigte Analysemerkmal „Fahrleistung in der aktuellen Periode“ weist für die Fahrzeuge in der HU-Stichprobe zu beinahe 100 % fehlende Werte<sup>22</sup> auf; zur Beseitigung dieser Datenlücken werden die mittleren bisherigen Fahrleistungen „imputiert“, d.h. den Fahrzeugen als aktuelle Fahrleistungswerte „zugeschrieben“.

## 6.4.2 Erste kritische Diskussion

Die endogene Fahrleistungsschätzung nach dem Durchschnittswertverfahren geht gemäß (6-6) von der Annahme aus, dass die durchschnittliche monatliche Fahrleistung eines Fahrzeugs im Bezugsjahr der Hochrechnung genauso groß ist wie die durchschnittliche monatliche Fahrleistung des betreffenden Fahrzeugs in seiner gesamten bisherigen Lebenszeit. Unter dieser Prämisse variiert die Monatsfahrleistung wohl von Fahrzeug zu Fahrzeug, für ein einzelnes Fahrzeug ist die

---

<sup>22</sup> Lediglich für die im Prüffjahr erstmals zugelassenen Fahrzeuge beinhaltet die HU-Stichprobe Angaben zur Fahrleistung in der aktuellen Periode.

Monatsfahrleistung aber im Zeitverlauf konstant<sup>23</sup>. Trifft diese Annahme zu, so wächst der Tachostand eines Fahrzeugs linear mit dem Fahrzeugalter.

Die Annahme, dass bei Fahrzeugen, die über die Zeit hinweg beobachtet werden, die Monatsfahrleistung konstant ist, scheint auf den ersten Blick im Widerspruch zu Ergebnissen der empirischen Verkehrsforschung zu stehen:

- Auswertungen der Fahrleistungserhebung 2002 (vgl. Hautzinger, Stock, Schmidt u.a., 2005) zeigen Folgendes: Gruppiert man die Fahrzeuge in der Stichprobe nach dem Fahrzeugalter, so unterscheidet sich die durchschnittliche tägliche Fahrleistung im Erhebungsjahr (also die mittlere aktuelle Tagesfahrleistung) zwischen den unteren Altersgruppen relativ wenig. Bei den oberen Altersgruppen stellt man dagegen eine mit wachsendem Fahrzeugalter sinkende durchschnittliche tägliche Fahrleistung im Erhebungsjahr fest.

Da es sich bei der Fahrleistungserhebung um eine Querschnittsstudie handelt, darf der empirische Befund, wonach die mittlere Tagesfahrleistung 2002 der älteren Fahrzeuge niedriger ist als die mittlere Tagesfahrleistung 2002 der jüngeren Fahrzeuge, nun aber auf keinen Fall dahingehend interpretiert werden, dass eine Beobachtung einzelner Fahrzeuge im Zeitverlauf<sup>24</sup> zu dem Ergebnis führen würde, dass im Zuge der Alterung der Fahrzeuge deren Tagesfahrleistung sinkt: Auch wenn bei jedem einzelnen Fahrzeug einer Kohorte (gleicher Zulassungsmonat) die Tagesfahrleistung zeitlich konstant ist und die *individuelle* Tachostandsfunktion damit *linear* wächst, wird man bei der *aggregierten* Tachostandsfunktion der Kohorte (Mittelwert der individuellen Tachostände über alle jeweils noch im Bestand befindlichen Fahrzeuge) ein im Zeitverlauf *degressives* Wachstum feststellen. Ursächlich hierfür ist, dass die stark genutzten Fahrzeuge aus der Kohorte relativ früh einen hohen Tachostand erreichen und deshalb früher als die weniger intensiv genutzten Fahrzeuge aus dem Bestand ausscheiden.

Für die Fahrleistung in der aktuellen Periode sind naturgemäß nur diejenigen Fahrzeuge einer Kohorte maßgeblich, welche die betreffende Periode tatsächlich „erleben“. Die obigen Ausführungen zeigen, dass bei den „überlebenden“ Fahrzeugen einer Kohorte sehr wohl die aktuelle mit der gesamten bisherigen Tagesfahrleistung übereinstimmen kann, auch wenn man bei Querschnittsanalysen eine Altersabhängigkeit der aktuellen Tagesfahrleistung feststellt!

Maßgeblich für die Validität des Modells M4 bzw. die Gültigkeit von Gleichung (6-6) ist, ob auf der *Individualebene* (einzelne Fahrzeuge einer Kohorte) die Tages- bzw. Monatsfahrleistung mit wach-

---

<sup>23</sup> Genauer gesagt wird unterstellt, dass die *erwartete* Monatsfahrleistung eines Fahrzeugs nicht von dem im betreffenden Monat erreichten Lebensalter des Fahrzeugs abhängt.

<sup>24</sup> also bei einer Längsschnittstudie



sendem Alter sinkt oder - wie im Modell M4 postuliert - konstant bleibt. Eine detaillierte empirische Überprüfung dieser für die endogene Schätzung der periodenbezogenen Fahrleistung zentralen Annahme ist derzeit leider nicht möglich, da hierfür Langzeitbeobachtungen des Tachostands einzelner Fahrzeuge nötig wären. Gleichwohl gibt es aber empirische Befunde, welche das Konzept der endogenen Schätzung stützen (vgl. Kapitel 7).

### **Ergänzende Anmerkungen**

Am nachfolgenden hypothetischen Beispiel soll ergänzend verdeutlicht werden, dass selbst dann, wenn bei einem Teil der Fahrzeuge die individuelle Tachostandsfunktion im Zeitablauf degressiv wachsen würde, man mit der Durchschnittswertmethode nicht zwangsläufig zu einer Überschätzung des Totalwerts der Fahrleistung in der Bezugsperiode der Hochrechnung käme.

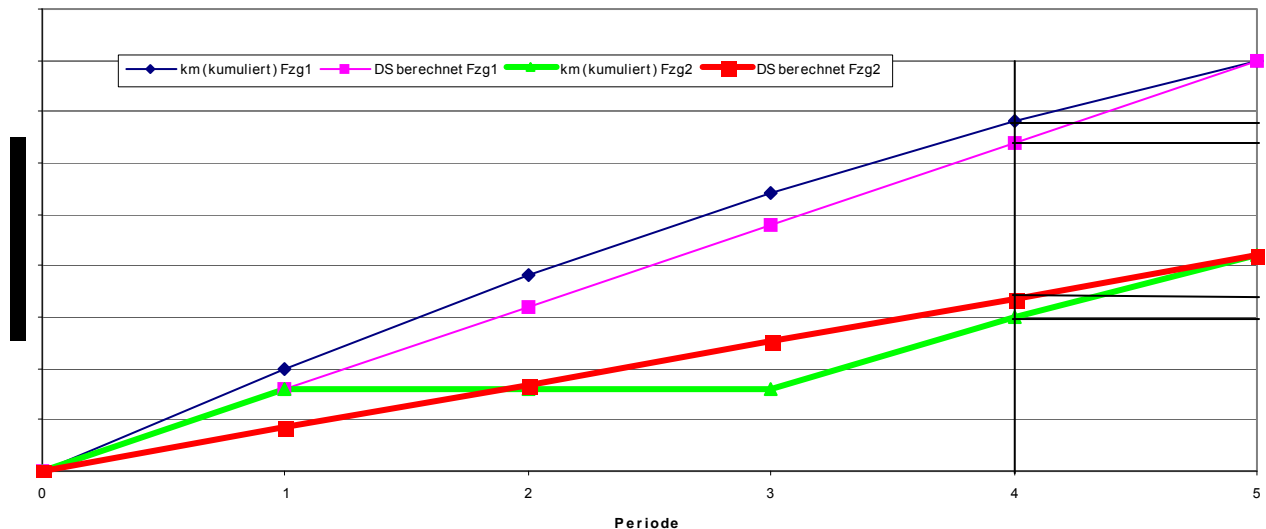
Naturgemäß nimmt der Tachostand eines Fahrzeugs im Zeitverlauf nur dann zu, wenn das Fahrzeug gefahren wird. Weist ein in der Bezugsperiode angemeldetes Fahrzeug im Verlauf seines bisherigen Lebens in größerem Umfang Zeiten vorübergehender Stilllegung auf, so kann die aktuelle Fahrleistung dieses Kraftfahrzeugs durchaus höher als seine mittlere bisherige Fahrleistung. Für Fahrzeuge, auf welche dies zutrifft, *unterschätzt* also die Durchschnittswertmethode die Fahrleistung in der Bezugsperiode. Die folgende Abbildung 3 verdeutlicht den geschilderten Sachverhalt. In dieser Abbildung stellt Zeitpunkt 5 das Prüfdatum dar, an welchem der Tachostand für Zwecke der Fahrleistungsschätzung abgelesen wird. Der Zeitraum zwischen den Zeitpunkten 4 und 5 ist die aktuelle Periode, auf welche sich die Hochrechnung bezieht.

Für Fahrzeug 1 („Fzg1“), bei welchem mit wachsendem Lebensalter die Fahrleistung je Periode abnimmt, würde man für die aktuelle Periode mit der Durchschnittswertmethode eine *zu hohe* Fahrleistung ermitteln, da die Steigung der blauen Linie (tatsächliche Entwicklung des Tachostands) zwischen den Zeitpunkten 4 und 5 niedriger ist als die Steigung der violetten Linie (unterstellte Entwicklung des Tachostands gemäß Durchschnittswertmethode).

Genau das Gegenteil gilt für Fahrzeug 2 („Fzg2“), welches zwischen Erstanmeldungsdatum und Prüfdatum über zwei Perioden hinweg stillgelegt war, in der aktuellen Periode zwischen den Zeitpunkten 4 und 5 aber angemeldet ist. Die bei Fahrzeug 2 für die aktuelle Periode unterstellte Fahrleistung (Anstieg der roten Linie) ist geringer als die tatsächliche Fahrleistung (Anstieg der grünen Linie). Für Fahrzeug 2 erhält man mit der Durchschnittswertmethode für die aktuelle Periode also eine *zu niedrige* Fahrleistung.

Das hypothetische Beispiel zeigt, dass bei Anwendung der Durchschnittswertmethode auf der Ebene einzelner Fahrzeuge Über- und Unterschätzungen der aktuellen Fahrleistung möglich sind, die in der Summe zu einem gewissen Ausgleich und damit zu einer hinreichend genauen Quantifizierung der Fahrleistungen in der aktuellen Periode führen können.

**Abbildung 3: Berechnung der Fahrleistung für die aktuelle Periode auf Einzeldatenbasis nach der Durchschnittswertmethode (hypothetisches Beispiel)**



Als Zwischenfazit lässt sich nach diesen Überlegungen festhalten, dass die Grundannahme der endogenen Fahrleistungsschätzung nach der Durchschnittswertmethode durchaus im Einklang mit vorliegenden Ergebnissen der empirischen Verkehrsforschung steht. Die bei Querschnittsuntersuchungen (Fahrleistungserhebung, KiD, MiD) übereinstimmend festgestellte Abnahme der Tagesfahrleistung mit dem Fahrzeugalter ist kein Beleg für mangelnde Realitätsnähe des Modells M4.

## 6.5 Resümee

In Abwägung der Eigenschaften der verschiedenen Methoden zur Schätzung kalenderjahrbezogener Fahrleistungen auf Basis von HU-Daten kann die Durchschnittswertmethode als die am ehesten geeignete beurteilt werden. Da die Durchschnittswertmethode die Periodenfahrleistung der einzelnen Altersklassen „endogen“ ermittelt, werden insbesondere die Probleme vermieden, die bei einer exogenen Schätzung (empirische Tachostandsdifferenzen zwischen Kohorten bzw. Perioden sowie Periodenfahrleistung als erste Ableitung der aggregierten Tachostandsfunktion) daraus resultieren, dass innerhalb einer Kohorte - also bei gegebenem Zeitpunkt der Erstzulassung - Fahrzeuge mit hoher Jahresfahrleistungen und damit schnell wachsendem Tachostand früher als andere aus dem Bestand ausscheiden.

Vor einer detaillierten Ausformulierung der Durchschnittswertmethode wird im nachfolgenden Kapitel 7 die Validität dieser Methode empirisch untersucht. Eine solche Validitätsprüfung ist mit Daten der bundesweiten Fahrleistungserhebung 2002 sowie der Erhebung von Fahrleistung und Kraftstoffverbrauch im Rahmen des Deutschen Mobilitätspanels möglich.

## 7 Empirische Prüfung des Konzepts der endogenen Schätzung

### 7.1 Überprüfung anhand der Fahrleistungserhebung 2002

Da bei der Fahrleistungserhebung 2002 (kurz FLE 2002) der Tachostand der in die Stichprobe gelangten Fahrzeuge wiederholt, nämlich im Abstand von 10 Wochen, abgelesen worden ist, kann dort auf Einzeldatenebene nicht nur die *mittlere bisherige Monatsfahrleistung*, resultierend aus dem aktuellen Tachostand und der Zahl der Tage seit Erstzulassung sondern auch die *aktuelle Monatsfahrleistung*, resultierend aus der Fahrleistung im 10-wöchigen Berichtszeitraum, berechnet werden. Durch Gegenüberstellung der Mittelwerte der beiden genannten Variablen lässt sich beurteilen, inwieweit die zentrale Annahme der Durchschnittswertmethode

„Bei Fahrzeugen, die das Untersuchungsjahr (hier 2002) erleben, ist die aktuelle Durchschnittsfahrleistung im Untersuchungsjahr näherungsweise genauso groß ist wie die Durchschnittsfahrleistung in der gesamten bisherigen Lebenszeit vor dem Untersuchungsjahr“

zutrifft. Die Ergebnisse des entsprechenden Vergleichs sind in der folgenden Tabelle für verschiedene Fahrzeuggruppen dargestellt.

**Tabelle 4: Aktuelle Monatsfahrleistung und mittlere bisherige Monatsfahrleistung gegliedert nach Fahrzeuggruppen (Fahrleistungserhebung 2002)**

Fahrzeuggruppe	Monatsfahrleistung	Mittelwert (km/Monat)	N	Standardfehler	Verhältnis aktuelle zu mittlere bisherige Fahrleistung
Krafträder	bisherig	265	2.196	4,54	121%
	aktuell	321	2.196	11,24	
Pkw	bisherig	1.296	20.330	4,26	103%
	aktuell	1.339	20.330	7,10	
Lkw	bisherig	2.046	13.679	20,06	105%
	aktuell	2.153	13.679	22,83	
Sattelzugmaschinen	bisherig	9.212	3.034	70,62	96%
	aktuell	8.888	3.034	80,42	
Sonst. Zugmaschinen	bisherig	812	1.586	25,06	78%
	aktuell	632	1.586	26,01	
Omnibusse	bisherig	4.666	2.438	34,12	94%
	aktuell	4.387	2.438	46,31	
Sonstige Kfz	bisherig	1.031	2.687	16,74	102%
	aktuell	1.054	2.687	25,38	

Wie die Abweichungskennzahlen in der rechten Spalte zeigen, weichen die beiden Fahrleistungsmittelwerte („aktuell“ versus „bisherig“) bei den bestands- und fahrleistungsmäßig bedeutenden Fahrzeuggruppen um nur wenige Prozentpunkte voneinander ab:

- Pkw (+3 %)
- Lkw (+5 %)
- Sattelzugmaschinen (-4 %)
- Omnibusse (-6 %)
- Sonstige Kfz (+2 %)

Die obigen fünf Fahrzeuggruppen mit guter Übereinstimmung von aktueller und bisheriger Durchschnittsfahrleistung machen zusammen über 90 % aller zugelassenen Fahrzeuge aus.

Deutliche Unterschiede zwischen aktueller und bisheriger Durchschnittsfahrleistung gibt es lediglich bei den Fahrzeuggruppen

- Krafträder (+21 %)
- Sonstige Zugmaschinen (-22 %).

Dass bei Krafträdern die aktuelle Fahrleistung so deutlich über der bisherigen Durchschnittsfahrleistung liegt, dürfte mit den bei dieser Fahrzeuggruppe besonders hohen Stilllegungszeiten zusammenhängen.

Es versteht sich, dass die obigen Abweichungskennzahlen als Korrekturfaktoren in ein auf der Durchschnittswertmethode beruhendes Hochrechnungsverfahren eingebracht werden können.

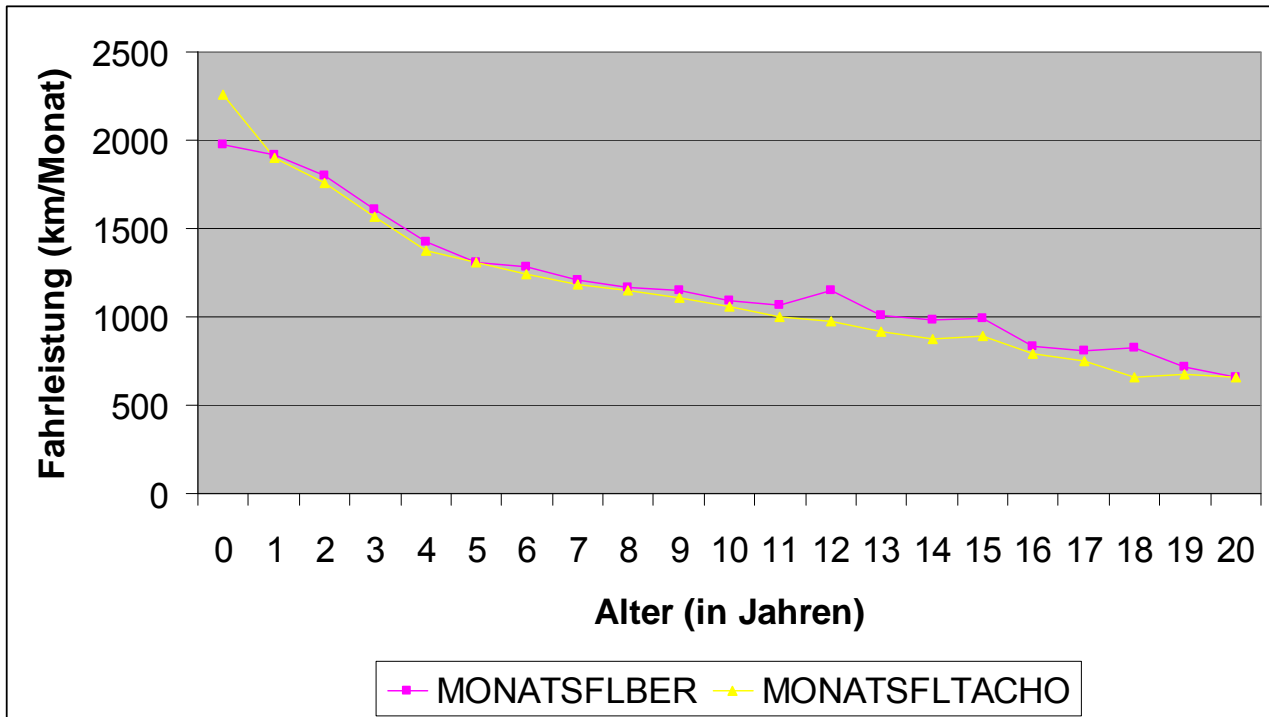
In der folgenden Tabelle ist für die Fahrzeuggruppe Pkw (private und gewerbliche Halter) die Abweichung zwischen aktueller und bisheriger Durchschnittsfahrleistung zusätzlich differenziert nach Zulassungskohorten dargestellt. Der Gesamtstichprobenumfang liegt hier bei ca. 20.000 Pkw

**Tabelle 5: Aktuelle Monatsfahrleistung und mittlere bisherige Monatsfahrleistung von Pkw gegliedert nach Jahr der Zulassung (Fahrleistungserhebung 2002)**

Jahr der Zulassung	Monatsfahrleistung	Mittelwert (km/Monat)	Anzahl Fälle	Standardfehler	Abweichungskennzahl
1982	bisherig	659	47	55,08	100,4%
	aktuell	662	47	102,32	
1983	bisherig	672	74	45,38	106,9%
	aktuell	719	74	93,54	
1984	bisherig	655	124	29,43	126,5%
	aktuell	828	124	78,98	
1985	bisherig	747	164	27,68	108,3%
	aktuell	809	164	74,13	
1986	bisherig	793	312	24,53	104,8%
	aktuell	831	312	57,00	
1987	bisherig	893	519	19,48	110,6%
	aktuell	988	519	53,64	
1988	bisherig	878	665	15,46	112,4%
	aktuell	987	665	41,32	
1989	bisherig	920	875	14,30	109,3%
	aktuell	1.006	875	34,82	
1990	bisherig	978	1.242	13,36	117,4%
	aktuell	1.148	1.242	35,97	
1991	bisherig	997	1.931	12,19	107,4%
	aktuell	1.070	1.931	25,19	
1992	bisherig	1.058	1.880	11,93	103,1%
	aktuell	1.091	1.880	23,87	
1993	bisherig	1.108	1.512	13,47	104,1%
	aktuell	1.154	1.512	27,61	
1994	bisherig	1.153	1.434	14,47	101,1%
	aktuell	1.165	1.434	26,73	
1995	bisherig	1.179	1.733	14,73	102,5%
	aktuell	1.209	1.733	25,92	
1996	bisherig	1.241	1.795	16,47	103,5%
	aktuell	1.285	1.795	28,66	
1997	bisherig	1.308	1.860	17,67	100,2%
	aktuell	1.310	1.860	27,72	
1998	bisherig	1.375	2.123	17,98	103,5%
	aktuell	1.424	2.123	27,75	
1999	bisherig	1.567	2.205	22,98	102,9%
	aktuell	1.612	2.205	30,30	
2000	bisherig	1.762	2.233	26,52	102,1%
	aktuell	1.799	2.233	33,45	
2001	bisherig	1.901	2.146	29,73	100,6%
	aktuell	1.913	2.146	35,13	
2002	bisherig	2.259	477	125,21	87,3%
	aktuell	1.973	477	81,65	

Die folgende Abbildung fasst die in der Tabelle dargestellten Ergebnisse zusammen.

**Abbildung 4: Aktuelle Monatsfahrleistung und mittlere bisherige Monatsfahrleistung von Pkw in Abhängigkeit des im Jahr 2002 erreichten Fahrzeugalters (Fahrleistungserhebung 2002)**



Legende: MONATSFLBER = aktuelle Monatsfahrleistung; MONATSFLTACHO = mittlere bisherige Monatsfahrleistung

Wie man sieht, stimmen für die stark besetzten Pkw-Altersgruppen (1 bis 10 Jahre) die aktuelle und bisherige Durchschnittsfahrleistung gut überein. In den oberen Altersgruppen liegt die aktuelle Monatsfahrleistung (entgegen mancher Erwartungen) tendenziell *über* der mittleren bisherigen Monatsfahrleistung. Die dargestellten empirischen Befunde stützen klar die Annahme, auf welcher die Durchschnittswertmethode beruht.

## 7.2 Überprüfung anhand der Erhebung von Fahrleistung und Kraftstoffverbrauch im Rahmen des Deutschen Mobilitätspanels

Im Rahmen des Deutschen Mobilitätspanels (MOP), einer seit 1994 jährlich in Form eines Rotationspanels durchgeführten Haushaltsbefragung zur Alltagsmobilität, werden in der sogenannten Erhebung von Fahrleistung und Kraftstoffverbrauch („Tankbuchehebung“) auch die Fahrleistungen von Pkw aus Privathaushalten erhoben. Jeweils im Frühjahr (etwa in den Monaten April bis Juni) wird rund die Hälfte der Haushalte, die an der Erhebung zur Alltagsmobilität teilnehmen, aufgefordert, Angaben zur Nutzung ihrer Pkw zu machen und insbesondere für jeden Tankvorgang im Berichtszeitraum das Datum, die getankte Menge und den aktuellen Tachostand des Fahrzeugs aufzuzeichnen.

Beispielsweise umfasst die Stichprobe der Tankbuchehebung 2007 (nach der Plausibilisierung) 423 Pkw privater Halter. Für diese Fahrzeuge wurden über einen Zeitraum von zwei Monaten die Tankvorgänge mit Datum und Kilometerstand erfasst. Da das Baujahr der Fahrzeuge bekannt ist, kann für jedes Fahrzeug in der Stichprobe eine hypothetische Jahresfahrleistung 2007 nach zwei verschiedenen Konzepten berechnet werden:

- Methode A: Schätzung der Fahrleistung 2007 mit „aktuellen“ Daten basierend auf der Länge des Berichtszeitraums und dem Tachostand zu Beginn und Ende des Berichtszeitraums (Umlegung der Fahrleistung im Berichtszeitraum auf das Bezugsjahr der Hochrechnung)
- Methode B: Schätzung der Fahrleistung 2007 nach der Durchschnittswertmethode, d.h. mit „historischen“ Daten, basierend auf dem Tachostand am Ende der Berichtsperiode und dem aus dem Baujahr berechneten Fahrzeugalter (Umlegung der gesamten bisherigen Fahrleistung auf das Bezugsjahr der Hochrechnung)

In der folgenden Tabelle sind die Ergebnisse der Schätzung der mittleren Jahresfahrleistung 2007 nach beiden Methoden dargestellt.

**Tabelle 6: Mittlere Fahrleistung 2007 von Pkw aus Privathaushalten gegliedert nach Alter berechnet auf Basis aktueller empirischer Fahrleistungsdaten sowie nach der Durchschnittswertmethode (Mobilitätspanel 2007)**

Alter des Fahrzeugs	Anzahl Pkw	Empirische aktuelle Fahrleistung (A) (km/Jahr)	Durchschnittswertmethode (B) (km/Jahr)	Verhältnis A : B in %
0 - 3 Jahre	109	14.472	13.949	103,8 %
4 - 6 Jahre	97	12.575	12.347	101,9 %
7 - 9 Jahre	86	12.923	12.926	100,0 %
10 Jahre und älter	131	13.846	14.495	95,5 %
Gesamt	423	13.528	13.543	99,9 %

Wie man sieht, liefern die Methoden A (aktuelle Fahrleistung) und B (Durchschnittswertmethode) für die mittlere Pkw-Jahresfahrleistung 2007 nahezu identische Ergebnisse:

- Methode A: 13.528 km/Jahr
- Methode B: 13.543 km/Jahr.

Gewisse Unterschiede zeigen sich bei einer Betrachtung nach Altersgruppen:

- Bei neuen bzw. jüngeren Fahrzeugen in der Altersgruppe 0 bis 3 Jahre liegt die aus der Fahrleistung am „aktuellen Rand“ berechnete Jahresfahrleistung 2007 um ca. 4 % *über* dem Wert, der sich ergibt, wenn man nach der Durchschnittswertmethode die mittlere Fahrleistung seit Erstzulassung auf das aktuelle Untersuchungsjahr 2007 überträgt.
- Bei Fahrzeugen im Alter zwischen 4 und 9 Jahren führen beide Berechnungsverfahren zu annähernd demselben Resultat.
- Für alte Fahrzeuge (10 Jahre und älter) kehrt sich dieser Befund gerade um. Hier liegt die aktuelle Fahrleistung um ca. 5 % *unter* dem Ergebnis gemäß Durchschnittswertmethode.

Insgesamt spricht auch dieses Resultat dafür, dass bei der Schätzung der periodenbezogenen Jahresfahrleistung von Pkw privater Halter eine Hochrechnung nach der Durchschnittswertmethode zumindest dann zu validen Ergebnissen führen kann, wenn bei der Ergebnisdarstellung über die Altersgruppen hinweg zusammengefasst wird<sup>25</sup>.

### **7.3 Zusammenfassende Bewertung der Ergebnisse der empirischen Überprüfung der Durchschnittswertmethode**

Zusammenfassend kann festgestellt werden, dass der Ergebnisvergleich durchaus dazu ermutigt, ein auf der Durchschnittswertmethode basierendes Hochrechnungsverfahren zur Schätzung der aktuellen Jahresfahrleistung von Fahrzeugen auf der Basis von Tachostandsdaten aus Hauptuntersuchungen

- im methodischen Detail auszuarbeiten und
- auf der Basis von konkreten HU-Daten zu erproben.

Für die Etablierung einer laufenden Fahrleistungsstatistik reicht das Vorhandensein eines Hochrechnungsverfahrens und einer zugehörigen Hochrechnungssoftware nicht aus. Notwendig ist vielmehr auch ein auf die Hochrechnungsmethodik abgestimmtes Verfahren zur Prüfung und Aufbereitung der Rohdaten, auf denen eine solche Statistik letztlich aufgebaut ist. Ein solches Verfahren wird im Folgenden dargestellt.

---

<sup>25</sup> Beim Vergleich der Prüfergebnisse anhand der FLE- und MOP-Daten ist zu berücksichtigen, dass die MOP-Daten sich nur auf Pkw privater Halter beziehen und bei MOP die Berechnung der bisherigen Lebensdauer nur grob (im Vergleich zur FLE 2002) erfolgen kann.



## **8 Verfahren zur Aufbereitung der HU-Rohdaten für Zwecke einer Fahrleistungsstatistik**

Zur Prüfung, Plausibilisierung und Zusammenführung der HU-Rohdaten eines bestimmten Bezugsjahres für Zwecke einer Fahrleistungsstatistik wurde im Rahmen des vorliegenden Projekts ein spezifisches Verfahren entwickelt und praktisch erprobt. Dieses Verfahren wird im Folgenden beschrieben und anhand der HU-Daten 2006 und 2007 illustriert.

Im Rahmen des Aufbereitungsverfahrens sind erste elementare Prüfungen und Bereinigungen bereits getrennt für die einzelnen Rohdatendateien durchzuführen. Der Großteil der Verfahrensschritte bezieht sich jedoch auf die Gesamtrohdatendatei des betreffenden Prüfjahres, die - ebenfalls im Rahmen der Datenaufbereitung - durch Zusammenführung der institutionsspezifischen Einzeldateien gebildet wird. Das vorliegende Kapitel beschreibt das methodische Vorgehen bei der Datenaufbereitung, eine Darstellung der programmtechnischen Umsetzung des Verfahrens mit Hilfe von SPSS findet sich in Band 2 des Schlussberichts.

### **8.1 Vereinheitlichung von Formaten**

Die von den teilnehmenden Prüforganisationen bereitgestellten Rohdatendateien können in unterschiedlichen Textformaten generiert worden sein; diesem Umstand ist bei der Zusammenführung und Aufbereitung der HU-Daten durch entsprechende Gestaltung der Einleseprozeduren für die einzelnen Rohdatendateien Rechnung zu tragen. Auf demselben Wege sind Datumsformate (Jahr und Monat der Erstzulassung bzw. Prüfung) zu vereinheitlichen, sofern in dieser Beziehung von den datenliefernden Institutionen unterschiedliche Codierungen verwendet werden.

### **8.2 Bereinigung um Dubletten und Fälle mit ungültigen Variablenwerten**

Im Rahmen der Prüfung der Datensätze der einzelnen Rohdatendateien sind im ersten Schritt „Doppelte Fälle“ zu eliminieren, wobei doppelte Fälle hier dadurch gekennzeichnet sind, dass die betreffenden Datensätze in den Variablen „Her“, „Typ“, „Monat“, „Jahr“ und „Tacho“ übereinstimmende Werte aufweisen. Der Anteil der Dubletten in den einzelnen Rohdatendateien beträgt 0,1 % bis 3,6 % des Gesamtumfangs.

Im Zuge der Datenbereinigung sind aus den einzelnen Rohdatendateien Datensätze (Fälle) mit ungültigen Codes bei den Variablen „Her“ und „Typ“ zu eliminieren; die Anzahl solcher unbrauchbarer Datensätze ist bei den HU-Daten 2006 und 2007 allerdings nur marginal.

Eine weitere Gültigkeitsprüfung muss sich auf die Codierung der Datumsangaben beziehen<sup>26</sup>. Bei den HU-Daten 2006 und 2007 wiesen alle Fälle gültige Datumsangaben auf.

Der Umfang der nach den eben beschriebenen Bereinigungsverfahren verbleibenden Rohdatendateien (jeder Datensatz in der Datei entspricht einer Fahrzeugprüfung) wird am Beispiel der HU-Daten 2007 in folgender Tabelle dargestellt:

**Tabelle 7: Anzahl verbleibender HU-Datensätze 2007 nach Bereinigung um Dubletten und Fällen mit ungültigen Codes gegliedert nach Prüforganisationen**

Datenliefernde Prüforganisation	Anzahl Rohdatensätze
Hessen TP	257.656
Hessen ÜO	296.181
TÜV Thüringen	0
TÜV Pfalz UO	87.200
TÜV Pfalz TP	104.605
FSP	473.324
TÜV Rheinland TP	516.406
TÜV Rheinland UO	736.605
GTÜ	3.259.610
TÜV Süd	3.627.310
Gesamt	9.358.897

Anmerkung: Die oben beschriebenen Bereinigungsverfahren sind programmtechnisch betrachtet Teil der Einleseprozeduren des SPSS-Programms zur laufenden Hochrechnung kalenderjahrbezogener Fahrleistungen auf der Basis von HU-Daten.

### **8.3 Zusammenführung der einzelnen Rohdatendateien und Bereinigung um Fälle mit ungültigen Angaben zum Tachostand**

In einem weiteren Schritt der Datenaufbereitung müssen die einzelnen Rohdatendateien der datenliefernden Prüforganisationen zu einer Gesamtrohdatendatei zusammengefügt werden.

---

<sup>26</sup> Die auf das Kernmerkmal „Tachostand“ bezogenen Plausibilitätsprüfungen werden in Abschnitt 8.6 gesondert dargestellt.

Nach dieser Zusammenführung sind vor den in den Abschnitten 8.4 bis 8.6 beschriebenen weiteren Aufbereitungsschritten (Ergänzung des Datensatzes um fahrzeugbezogene Merkmale und Eliminierung von Fällen mit unplausiblen Angaben zum Tachostand) Datensätze mit ungültiger Angabe zum Tachostand zu eliminieren. Als ungültige Angabe wird ein negativer Kilometerstand oder der Kilometerstand Null betrachtet. In der Gesamtrohdatendatei für das HU-Prüfjahr 2007 weisen 12.830 Fälle eine ungültige Angabe zum Tachostand auf, was bei einem Ausgangswert von insgesamt 9.358.897 Datensätzen zu vernachlässigen ist. Damit stehen nach diesem Schritt der Datenaufbereitung für das Jahr 2007 von den insgesamt neun datenliefernden Prüforganisationen 9.346.067 Datensätze zu geprüften Fahrzeugen mit einem Kilometerstand größer als Null zur Verfügung.

## **8.4 Ergänzung der HU-Rohdaten um fahrzeugtechnische Merkmale**

### **8.4.1 Merkmal „Fahrzeugart“**

Sowohl im Hinblick auf das Hochrechnungsverfahren als auch mit Blick auf das Tabellenprogramm einer Fahrleistungsstatistik wird das Merkmal „Fahrzeugart“ als Gliederungsmerkmal benötigt. Da dieses Merkmal im HU-Datensatz so nicht enthalten ist, muss eine entsprechende neue Variable aus den vorhandenen Angaben zum Hersteller- und Typschlüssel gebildet werden. Vor dem Hintergrund der Fahrleistungserhebung 2002 und unter Berücksichtigung von Gegebenheiten und Erfordernissen im Bereich der amtlichen Verkehrsstatistik wurden für die neu zu bildende Variable Fahrzeugart folgende Ausprägungen vorgegeben:

02	Krafträder
03	Pkw
04	Lkw
06	Sattelzugmaschinen
07	Sonstige Zugmaschinen
08	Zugmaschinen LoF
09	Omnibusse
10	Sonstige Kfz

Das vergleichsweise aufwendige methodische Vorgehen zur Ermittlung der Fahrzeugart für jedes geprüfte Fahrzeug in der HU-Rohdatendatei wird im Folgenden ausführlich beschrieben. Den Ausgangspunkt bilden dabei, wie erwähnt, die beiden originären Variablen Hersteller- und Typschlüssel (Her bzw. Typ).

Beim Kraftfahrt-Bundesamt (KBA) liegt eine Referenzdatei vor, in der (fast) jeder Hersteller-Typ-Kombination die zugehörige Fahrzeugklasse und Aufbauart zugeordnet ist; zudem werden dort für jede aufgeführte Kombination weitere Merkmale festgehalten, wie z.B. max. Hubraum, max. zuläs-

siges Gesamtgewicht, Kraftstoffart, KW-Leistung, Jahr der Genehmigung usw.. Die genannte Referenzdatei, die vom KBA sukzessiv fortgeschrieben wird, enthält zurzeit 56.322 verschiedene Kombinationen von Hersteller- und Typ-Codes.

Liegen für ein Fahrzeug - ausgehend von Hersteller- und Typschlüssel - die beiden Merkmale Fahrzeugklasse und Aufbauart fest, so kann das betreffende Fahrzeug eindeutig einer der oben aufgeführten Fahrzeugarten zugeordnet werden<sup>27</sup>. Für diese Zuordnung kommen allerdings drei unterschiedliche Systematiken infrage: Zum einen die neue (und in Zukunft verbindliche) EU-Systematik, zum zweiten die nationale Systematik und zum dritten eine auslaufende (nationale) Systematik. Letztere ist Vorgänger der nationalen Systematik. Gemäß der zum Zeitpunkt der Registrierung eines Fahrzeugs gültigen Systematik erfolgt die Eintragung im Register; insoweit sind also alle drei Systematiken zu berücksichtigen, da – vereinfachend ausgedrückt – das Alter eines Fahrzeugs die maßgebliche Systematik und damit die spezifische Zuordnung bestimmt. Es müssten demnach also bei der Bildung der neuen Variablen „Fahrzeugart“ alle drei Systematiken berücksichtigt werden, was jedoch Probleme hervorrufen würde, da gleiche Codierungen in den beiden nationalen Systematiken unterschiedlichen Fahrzeugarten (gemäß der Beschreibung in der MOWIVE-Referenzdatei) entsprechen bzw. nicht immer disjunkte Abgrenzungen bestehen.

Es wurde daher entschieden, dass sich die Bildung der Variablen Fahrzeugart an der neuen EU-Systematik orientiert und eventuell „falsche“ Zuordnungen (im Sinn der alten Systematiken) in Kauf genommen werden können, da diese aufgrund der Kohortenentwicklung im Fahrzeugbestand immer weiter an Bedeutung verlieren. Damit werden z.B. Wohnmobile den Pkw zugeordnet und bestimmte Fahrzeuge mit speziellen Nutzungsarten wie z.B. Rettungsfahrzeuge und Feuerwehrfahrzeuge der Kategorie „Sonstige Kfz“.

Vom Kraftfahrt-Bundesamt wurde im Zusammenhang mit der Ergänzung der HU-Rohdaten um fahrzeugtechnische Merkmale der Kfz-Bestand zum Stichtag 01.01.2008 nach Hersteller (Her) und Typ (Typ) ausgezählt. Dabei zeigte sich, dass im ZFZR-Bestand insgesamt 51.010 Her-Typ-Kombinationen vorkommen. Diese Anzahl unterschiedlicher Her-Typ-Kombinationen ist nicht deckungsgleich mit der entsprechenden Anzahl aus der Her-Typ-Referenzdatei, da diese Referenzdatei zum einen Kombinationen enthält, die nicht mehr im ZFZR-Bestand zu finden sind und zum anderen gewisse Her-Typ-Kombinationen mit nur wenigen Fahrzeugen nicht in die Referenzdatei aufgenommen wurden, obwohl die betreffenden Kombinationen im ZFZR-Bestand vorkommen.

---

<sup>27</sup> Referenzdatei „MOWIVE-Kodierung-Fz-Systematik“ (vgl. Band 2, Kapitel 1)

Somit konnte davon ausgegangen werden, dass sich nicht alle in den HU-Rohdaten tatsächlich vorkommenden Her-Typ-Kombinationen in der Referenzdatei des Kraftfahrt-Bundesamtes finden. Es wurden daher solche Hersteller identifiziert, die in der Her-Typ-Referenzdatei nur mit einer einzigen Fahrzeugart vertreten sind. Die betreffenden Hersteller wurden in einer separaten Datei abgelegt<sup>28</sup>. Für HU-Rohdatensätze ohne Entsprechung in der Her-Typ-Referenzdatei, bei denen aber der Hersteller in der genannten separaten Datei vertreten ist, ist so zumindest eine Zuordnung zu Fahrzeugarten möglich (weitere Fahrzeugmerkmale können in diesem Fall allerdings nicht ermittelt werden).

Mittels dieser Vorgehensweisen wurde schließlich eine Datei erzeugt, die – soweit möglich – für jede in der Fahrzeugbestands- oder Referenzdatei auftretende Her-Typ-Kombination die zugehörige Fahrzeugart (inkl. weiterer Fahrzeugmerkmale) enthält. Insgesamt umfasst diese Datei 58.660 verschiedene Her-Typ-Kombinationen, wobei für 48.590 Kombinationen mindestens die Fahrzeugart bestimmt werden kann.

Die Anwendung des beschriebenen Verfahrens zur Bildung der neuen Variablen „Fahrzeugart“ auf die HU-Rohdaten 2007 lieferte folgendes Ergebnis:

**Tabelle 8: Verteilung der Fahrzeuge in der HU-Rohdatendatei 2007 nach Fahrzeugarten**

Fahrzeugart	Häufigkeit	Prozent
Krafträder	646807	7.0 %
Pkw	7569690	81.4 %
Lkw	486216	5.3 %
Sattelzugmaschinen	38215	0.4 %
Sonstige Zugmaschinen	16668	0.2 %
Zugmaschinen LoF	11063	0.1 %
Omnibusse	13628	0.1 %
Sonstige Kfz	29798	0.3 %
Keine Zuordnung	491017	5.3 %
Gesamt	9303102	100.0 %

#### 8.4.2 Merkmal „Antriebsart“

Als weiteres fahrzeugtechnisches Merkmal kann die Antriebsart nach dem beschriebenen Verfahren über die Her-Typ-Referenzdatei den HU-Rohdaten hinzugefügt werden. Wie Tabelle 9 zeigt,

<sup>28</sup> Vgl. Band 2 Kapitel 1 SPSS-Datendatei *HER-FzArt\_eindeutig\_Fehlend.sav*

gibt es unter den 9.303.102 Datensätzen der HU-Rohdatendatei 2007 nur 607.620 Datensätze (d.h. 6,5 %), bei denen dies nicht möglich ist.

**Tabelle 9: Verteilung der Fahrzeuge in der HU-Rohdatendatei 2007 nach Antriebsarten**

Antriebsart	Häufigkeit	Prozent
Benzin	6462645	69.5 %
Diesel	2165241	23.3 %
Hybrid	873	0.0 %
Sonstiger altern. Antrieb	66492	0.7 %
Unbekannt	231	0.0 %
Keine Zuordnung	607620	6.5 %
Gesamt	9303102	100.0 %

Die Verteilung nach Fahrzeugarten zeigt, dass hier bei 491.017 Datensätzen keine Zuordnung möglich ist. Die Differenz  $607.620 - 491.017 = 116.603$  entspricht der Zahl der Datensätze, die zwar um das Merkmal Fahrzeugart nicht aber um das Merkmal Antriebsart ergänzt werden können.

### 8.4.3 Leistungs- und Größenmerkmale

Über die Her-Typ-Referenzdatei können den HU-Rohdaten weitere fahrzeugtechnische Merkmale hinzugespielt werden. Im Rahmen des vorliegenden Entwicklungsprojekts waren dies die Merkmale „Motorleistung in KW“ (für die Fahrzeugarten Krafträder und Pkw) und „zulässiges Gesamtgewicht“ (für die Fahrzeugart Lkw). Näheres siehe Band 2, Kapitel 2.

### 8.4.4 Grenzen der Datenergänzung

Die Tatsache, dass nicht alle Datensätze der HU-Rohdatendatei mittels der Her-Typ-Referenzdatei um weitere Fahrzeugmerkmale ergänzt werden können, erklärt sich hauptsächlich durch sogenannte „ungetypte“ Fahrzeuge, d.h. Fahrzeuge, bei denen der Typ-Code „000“ vorliegt. Unter den 607.620 HU-Datensätzen des Jahres 2007 ohne Möglichkeit einer Ergänzung um fahrzeugtechnische Merkmale finden sich 595.797 Datensätze (98 %) mit dem Typ-Code „000“.

Anmerkung: Die Generierung des HU-Rohdatenbestandes für das Jahres 2006 erfolgte in modifizierter Form, da die Datenbestände 2006 der Prüforganisationen 1-9 bereits in Form eines SPSS-Datenfiles vorlagen. Diesem Datenfile wurden die Daten des TÜV Süd hinzugefügt, die als ASCII-Daten eingelesen wurden. In der programmtechnischen Verfahrensbeschreibung ist dies im Einzelnen dokumentiert (vgl. Band 2, Kapitel 3).

## **8.5 Ergänzung der HU-Rohdaten um das Merkmal Fahrzeualter**

Die Datensätze der HU-Rohdatendatei enthalten als originäre Variable Monat und Jahr der Erstzulassung sowie Monat und Jahr der Hauptuntersuchung. Aus diesen Variablen können zwei neue Variable zur Kennzeichnung des Fahrzeualters gebildet werden, die für die Beurteilung der Repräsentativität der im Rahmen des Datenlieferungsabkommens übermittelten HU-Daten für alle bei Hauptuntersuchungen anfallenden Daten sowie für die Hochrechnung der periodenbezogenen Fahrleistung mit Hilfe der stichtagsbezogenen Tachostandsdaten von Bedeutung sind.

### **8.5.1 Fahrzeualter zum Zeitpunkt der Hauptuntersuchung**

Im Rahmen der Statistik der Fahrzeuguntersuchungen (sog. FU-Statistik) übermitteln die technischen Prüforganisationen dem KBA im Sinne einer Totalerhebung Informationen über die durchgeführten Hauptuntersuchungen von Kraftfahrzeugen untergliedert nach dem Alter des Fahrzeugs zum Zeitpunkt der Prüfung, welches sich aus dem Prüfdatum und dem Datum der Erstzulassung berechnet. Das Alter der geprüften Fahrzeuge wird dabei in kategorisierter Form angegeben:

- bis einschließlich 3 Jahre
- über 3 Jahre bis einschließlich 5 Jahre
- über 5 Jahre bis einschließlich 7 Jahre
- über 7 Jahre bis einschließlich 9 Jahre
- über 9 Jahre.

Um die Repräsentativität der vorliegenden HU-Rohdaten für die Gesamtheit aller in der Bundesrepublik Deutschland durchgeführten Hauptuntersuchungen hinsichtlich des Fahrzeualters beurteilen zu können, wurde in den HU-Rohdatenbestand eine Variable „Prüfalter in Jahren“ eingeführt, die mit den in der FU-Statistik verwendeten Alterskategorien kompatibel ist.

Grundsätzlich lässt sich das Prüfalter in Jahren auf unterschiedliche Art und Weise definieren bzw. berechnen. Die HU-Rohdaten ermöglichen eine Definition anhand des im Monat der Hauptunter-

suchung erreichten Lebensalters des Fahrzeugs in Monaten (Differenz zwischen Monat des HU-Prüftermins und Monat der Erstanmeldung). Nach verschiedenen Sensitivitätsanalysen wurde entschieden, für die Zwecke des vorliegenden Projekts das Prüfalter des Fahrzeugs in Jahren analog zum Lebensalter eines Menschen zu definieren (erst mit der Vollendung eines vollen Jahres beginnt ein neues Lebensjahr):

Prüfalter in Jahren	Erreichtes Lebensalter zum Prüfzeitpunkt (in Monaten)
0 Jahre	unter 12 Monate
1 Jahr	12 bis unter 24 Monate
2 Jahre	24 bis unter 36 Monate
etc.	

Die so definierte Variable „Prüfalter in Jahren“ kann nach entsprechender Klassenbildung für Tabellierungen der verfügbaren HU-Daten zum Zweck eines Vergleichs mit der FU-Statistik verwendet werden.

### 8.5.2 Fahrzeugalter im Bezugsjahr der Hochrechnung

Im Hinblick auf die Zielsetzung des Projekts sind Altersklassen, die sich auf das Alter eines Kfz zum Prüfstermin beziehen, als Schichtungsmerkmal für die Hochrechnung ungeeignet, da man für Fahrzeuge ohne Hauptuntersuchung im Bezugsjahr der Hochrechnung logischerweise ein solches Alter nicht berechnen kann. Daher wurde für Hochrechnungszwecke (Bildung von Hochrechnungsgruppen) das „Fahrzeugalter im Bezugsjahr der Hochrechnung in Jahren“ aus dem Zulassungsjahr und dem Bezugsjahr der Hochrechnung abgeleitet (Differenz zwischen Bezugsjahr und Zulassungsjahr).

Am Beispiel des Bezugsjahres 2007 bedeutet dies Folgendes:

Zulassungsjahr des Fahrzeugs	Fahrzeugalter im Bezugsjahr 2007
2007	0 Jahre
2006	1 Jahr
2005	2 Jahre
etc.	



## 8.6 Plausibilitätsprüfungen bezogen auf das Merkmal Tachostand

Zentrales Analysemerkmal der HU-Daten ist der Tachostand der geprüften Fahrzeuge. Wegen der starken Streuung von Kfz-Fahrleistungen und der Empfindlichkeit von Fahrleistungsmittelwerten gegenüber „Ausreißern“ in den Daten, kommt der Prüfung der Plausibilität von Tachoständen erhebliche Bedeutung zu. Die Plausibilitätsprüfung bezieht sich nach dem hier entwickelten Verfahren auf die

- Unter- bzw. Überschreitung gewisser „Abschneidegrenzen“ sowie auf die
- Problematik der sog. „Tachoüberläufe“.

Datensätze, bei denen der Tachostand als unplausibel einzustufen ist, sind aus dem Datenbestand zu entfernen, um Verzerrungen der Hochrechnungsergebnisse zu vermeiden.

### 8.6.1 Abschneidegrenzen für Tachostände

Wesentliche Ursachen für fehlerhafte Angaben zum Merkmal Tachostand sind Fehler beim Ablezen des Tachostands sowie Fehler bei der Dateneingabe. Ob solche Fehler vorliegen, kann im Nachhinein natürlich nicht mehr ermittelt werden. Möglich ist lediglich eine Prüfung des angegebenen Tachostands auf Plausibilität.

Als unplausibel sind auf jeden Fall solche Tachostände zu betrachten, die unter Berücksichtigung des Fahrzeugalters als extrem hoch bzw. extrem niedrig einzustufen sind. So kommt es vor, dass der angegebene Tachostand einer durchschnittlichen bisherigen Monatsfahrleistung von 50.000 km und mehr entspricht (was eine Jahresfahrleistung von 600.000 km und mehr bedeuten würde). Im anderen Extrem berechnet sich eine mittlere bisherige Monatsfahrleistung von weniger als 1 km. Da aber auch weniger extreme Werte vorkommen, die gleichwohl als unplausibel zu betrachten sind, muss im Rahmen des hier zu entwickelnden Verfahrens zur Datenbereinigung - differenziert nach Fahrzeugart und Fahrzeugalter (Alter zum Zeitpunkt der Hauptuntersuchung) - eine untere und obere Abschneidegrenze für Tachostände festgelegt werden.

Um empirische Grundlagen für die Festlegung von Abschneidegrenzen zu schaffen, wurden die Tachostände hinsichtlich der Häufigkeit ihres Auftretens ausgezählt. Dabei zeigte sich, dass die Kilometerstände „1 km“ und „9.999.999 km“ relativ häufig auftreten, so dass im Rahmen des Bereinigungsverfahrens davon ausgegangen wird, dass diese Codes von den Prüforganisationen zur Kennzeichnung fehlender Werte („keine Angaben zum Kilometerstand“) eingetragen wurden. Weitere Analysen führten zu den nachfolgend dargestellten Festlegungen.

### 8.6.1.1 Untere Grenze

Bei fast allen Fahrzeugarten tritt bei „neuen“ Fahrzeugen mit einem Alter von höchstens 12 Monaten relativ häufig ein niedriger Tachostand und damit eine niedrige mittlere bisherige Fahrleistung auf (im Vergleich zu älteren Fahrzeugen). Da für diese Fahrzeuge nicht auszuschließen ist, dass sie nach ihrer Erstzulassung während eines längeren Zeitraums vorübergehend stillgelegt bzw. nicht in Betrieb waren, wurde bei allen Fahrzeugen mit einem Alter bis 12 Monate als untere Abschneidegrenze einheitlich ein Tachostand von 5 km festgelegt.

Ab einem Fahrzeugalter von 13 Monaten wurden je Fahrzeugart Mindestkilometerleistungen je Lebensdauermonat unterstellt, woraus eine lineare Funktion resultiert, welche die für die Plausibilitätsprüfung maßgeblichen Mindesttachostände in Abhängigkeit des Fahrzeugalters (in Monaten) darstellt. Für die Fahrzeugarten „Pkw“, „Lkw“, „Sattelzugmaschinen“ und „Kraftomnibusse“ wurde als minimale durchschnittliche Monatsfahrleistung 50 km festgelegt, so dass bei diesen Fahrzeugarten z.B. bei einem 10 Jahre alten Fahrzeug (Alter 120 Monate) ein Mindesttachostand von 5.400 km ( $= [120-12] \cdot 50$  km) gegeben sein muss. (Die Mindestfahrleistung von 5 km in den ersten 12 Monaten nach der Erstanmeldung wird hier der Einfachheit halber ignoriert). Für die übrigen Fahrzeugarten gehen die minimalen durchschnittlichen Monatsfahrleistungen aus der rechten Spalte von Tabelle 10 hervor.

Bezeichnet man mit  $y$  den Mindesttachostand (untere Abschneidegrenze) und mit  $x$  das Fahrzeugalter in Monaten, so lassen sich die genannten funktionalen Zusammenhänge differenziert nach der Fahrzeugart wie in der nachfolgenden Tabelle beschrieben darstellen.

**Tabelle 10: Untere Abschneidegrenze des Tachostands als Funktion des Fahrzeugalters differenziert nach Fahrzeugarten**

Fahrzeugart	Untere Abschneidegrenze ( $y$ ) als Funktion des Fahrzeugalters ( $x$ ) zum Zeitpunkt der Hauptuntersuchung	
	Fahrzeugalter ( $x$ ) bis 12 Monate	Fahrzeugalter ( $x$ ) ab 13 Monate
Pkw, Lkw, Sattelzugmaschinen, Omnibusse	$y=5$ [km]	$y=(x-12) \cdot 50$ [km]
Krafträder	$y=5$ [km]	$y=(x-12) \cdot 7$ [km]
Zugmaschinen LoF	$y=5$ [km]	$y=(x-12) \cdot 2$ [km]
Sonstige Zugmaschinen	$y=5$ [km]	$y=(x-12) \cdot 5$ [km]
Sonstige Kfz	$y=5$ [km]	$y=(x-12) \cdot 10$ [km]

### 8.6.1.2 Obere Grenze

Im HU-Datenbestand finden sich auch extrem hohe Werte des Tachostands, sodass als weiteres Element des Verfahrens der Datenbereinigung eine obere Abschneidegrenze festgelegt werden muss. In diesem Zusammenhang wurde vorab entschieden, dass analog zum Wert „1 km“ der ebenfalls auftretende Kilometerstand „9.999.999 km“ als fehlender Wert zu behandeln ist.

Die oberen Grenzwerte für plausible Tachostände wurden getrennt nach Fahrzeugarten für die drei Fahrzeugaltersklassen

- bis 10 Jahre (120 Monate)
- von 11 bis 20 Jahre (121 bis 240 Monate)
- ab 21 Jahre (ab 241 Monate)

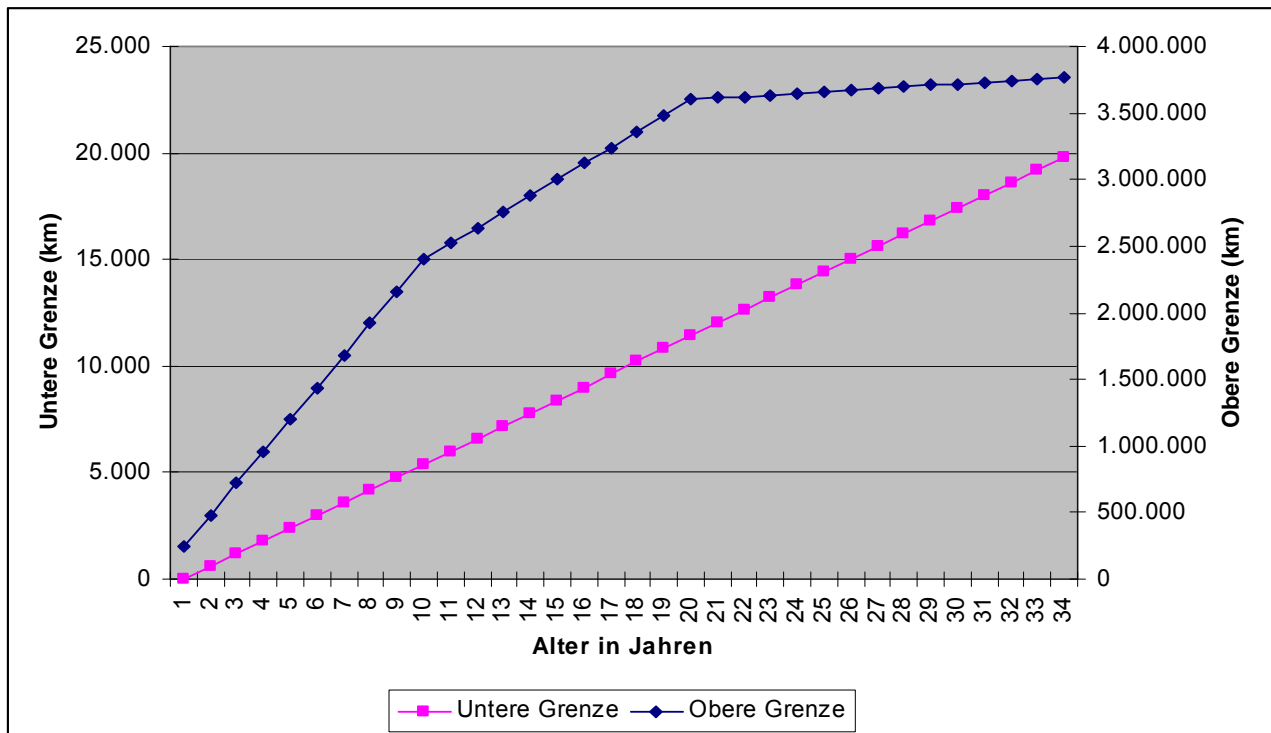
mittels (abschnittsweise) linearer Funktionen festgelegt. In Tabelle 11 ist dargestellt, wie für die einzelnen Fahrzeugarten die maximal zulässigen Tachostände (y) aus dem Fahrzeugalter in Monaten (x) zu berechnen sind. Generell wird für Zwecke der Plausibilitätsprüfung und Datenbereinigung von einem degressiven Wachstum der oberen Abschneidegrenze ausgegangen.

**Tabelle 11: Obere Abschneidegrenze des Tachostands als Funktion des Fahrzeugalters differenziert nach Fahrzeugarten**

Fahrzeugart	Fahrzeugalter		
	bis 10 Jahre	11 bis 20 Jahre	ab 21 Jahre
Pkw	$y=x*20.000$ [km]	$y=120*20.000+(x-120)*10.000$ [km]	$y=120*20.000+120*10.000+(x-240)*1.000$ [km]
Lkw, Sattelzugmaschinen, Omnibusse	$y=x*30.000$ [km]	$y=120*30.000+(x-120)*10.000$ [km]	$y=120*30.000+120*10.000+(x-240)*1.000$ [km]
Krafträder	$y=x*4.000$ [km]	$y=120*4.000+(x-120)*2.000$ [km]	$y=120*4.000+120*2.000+(x-240)*1.000$ [km]
Sonstige Zugmaschinen	$y=x*5.000$ [km]	$y=120*5.000+(x-120)*2.500$ [km]	$y=120*5.000+120*2.500+(x-240)*1.250$ [km]
Sonstige Kfz	$y=x*7.000$ [km]	$y=120*7.000+(x-120)*3.500$ [km]	$y=120*7.000+120*3.500+(x-240)*350$ [km]
Zugmaschinen LoF	$y=x*1.000$ [km]	$y=120*1.000+(x-120)*500$ [km]	$y=120*1.000+120*500+(x-240)*500$ [km]

In der folgenden Abbildung ist der funktionale Verlauf der unteren und oberen Abschneidegrenze für die Fahrzeugart Pkw dargestellt, wobei auf die unterschiedlichen Ordinatenkalierungen hinzuweisen ist.

**Abbildung 5: Untere und obere Abschneidegrenze des Tachostands als Funktion des Fahrzeugalters für die Fahrzeugart Pkw**

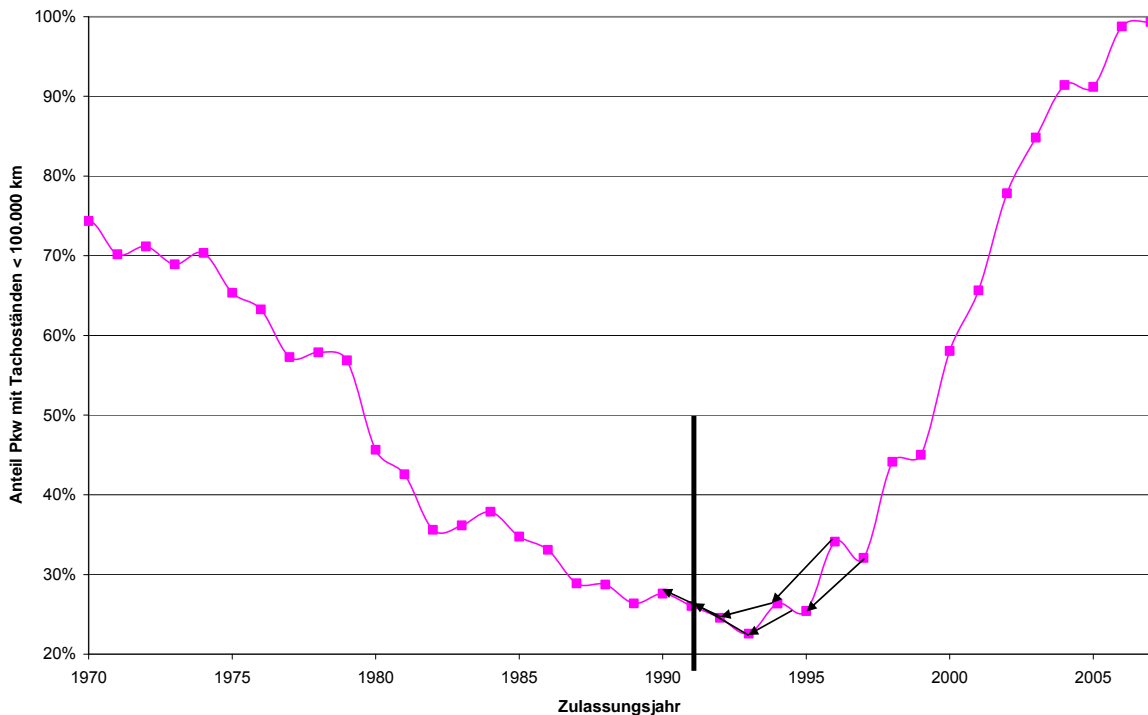


## 8.6.2 Korrektur von Verzerrungen durch Tachoüberläufe

Bei älteren Fahrzeugen mit entsprechend hoher Fahrleistung kann es durchaus zu Tachoüberläufen kommen, vor allem wenn der Wegstreckenzähler nur fünf Stellen ausweist. Als Folge eines oder mehrerer Tachoüberläufe ist der vom 5-stelligen Wegstreckenzähler eines Fahrzeugs angezeigte Kilometerstand gravierend niedriger (Vielfaches von 100.000 km) als die Zahl der tatsächlich gefahrenen Kilometer. Da 5-stellige Wegstreckenzähler nur bei älteren Baujahren vorkommen, zeitigt aufgrund der Kohortenentwicklung der Fahrzeugflotte dieses Problem in Zukunft glücklicherweise immer geringere negative Auswirkungen.

Auf der Basis der HU-Rohdaten wurden Analysen durchgeführt, um zu ermitteln, vor welchem Zulassungsjahr Tachoüberläufe ein Problem für die Hochrechnungsgenauigkeit darstellen können. Dazu wurde je Erstzulassungsjahr der Anteil von Fahrzeugen bestimmt, die einen Tachostand von *weniger* als 100.000 km aufweisen. Aufgrund einfacher Plausibilitätsüberlegungen müsste dieser Anteil mit zunehmendem Fahrzeugalter abnehmen. Wie die folgende Abbildung verdeutlicht, sinkt bei der Fahrzeugart Pkw der betreffende Anteil ausgehend vom Erstzulassungsjahr 2007 bis zum Erstzulassungsjahr 1991 tatsächlich stetig ab, steigt dann aber mit weiter wachsendem Fahrzeugalter wieder an. Bei Pkw, die vor dem Jahr 1991 zugelassen worden sind, kann demnach davon ausgegangen werden, dass Tachoüberläufe vorkommen.

**Abbildung 6: Pkw mit Tachostand unter 100.000 km in v.H. aller Pkw gegliedert nach Jahr der Erstzulassung (Basis: HU-Daten 2007)**



Auf Grundlage der skizzierten Analysen wurde für jede Fahrzeugart ein für die Korrektur von Verzerrungen durch Tachoüberläufe maßgebliches Erstzulassungsjahr bestimmt (Tabelle 12)

**Tabelle 12: Für die Tachostandskorrektur maßgebliches Erstzulassungsjahr**

Fahrzeugart	Erstzulassungsjahr
Pkw	1991
Lkw	1990
Sattelzugmaschinen	1990
Omnibusse	1991
Krafträder	1984
Sonstige Zugmaschinen	1990
Sonstige Kfz	1990
Zugmaschinen LoF	-

Nach dem hier entwickelten Bereinigungsverfahren erfolgt eine Korrektur des Tachostands von Fahrzeugen, bei denen aufgrund ihres Erstzulassungsjahres Tachoüberläufe möglich sind, dann, wenn der angegebene Tachostand die für die betreffende Fahrzeugart maßgebliche untere Abschneidegrenze unterschreitet.

Konkret wird für die entsprechenden Fahrzeuge<sup>29</sup> der in den HU-Daten angegebene Tachostand nach der folgenden Formel korrigiert:

$$\begin{aligned} &\text{Korrigierter Tachostand} \\ &= \text{angegebener Tachostand} + \\ &\quad + (\text{Untersuchungsjahr} - \text{Zulassungsjahr}) \times \text{hypothetische Jahresfahrleistung} \end{aligned}$$

Die bei der Korrektur zu verwendende „hypothetische Jahresfahrleistung“ geht aus der nachfolgenden Tabelle hervor.

**Tabelle 13: Hypothetische Jahresfahrleistungen zur Korrektur von Tachoüberläufen**

Fahrzeugart	Hypothetische Jahresfahrleistung [km/Jahr]
Pkw	12.500
Lkw	20.000
Sattelzugmaschinen	20.000
Omnibusse	20.000
Krafträder	6.000
Sonstige Zugmaschinen	10.000
Sonstige Kfz	12.500
Zugmaschinen LoF	-

Insgesamt ist das hier entwickelte Verfahren zur Behandlung des Problems der Tachoüberläufe nicht als Konzept zur Verbesserung der Genauigkeit individueller Tachostandsangaben zu betrachten sondern vielmehr als Konzept zur Vermeidung oder Reduzierung einer systematischen Unterschätzung der periodenbezogenen Gesamtfahrleistung bei den oberen Fahrzeugaltersklassen: Es versteht sich, dass auch bei Fahrzeugen, deren Tachostand über der unteren Grenze liegt, ein Tachoüberlauf vorliegen kann und dass nicht bei jedem Fahrzeug mit „zu geringer“ Fahrleistung ein Tachoüberlauf hierfür die Ursache ist.

Mit Blick auf die Hochrechnungsgenauigkeit ist das dargestellte Problem im Übrigen bereits heute kaum noch von Bedeutung; der durchschnittliche Anteil der Fahrzeuge in den HU-Daten, denen nach dem oben beschriebenen Verfahren ein korrigierter Tachostand zugeordnet wird, beträgt lediglich 0,3 %.

---

<sup>29</sup> Diese Korrektur bezieht sich nur auf Fahrzeuge mit einem Tachostand größer als Null, da Fälle mit negativem Tachostand oder Tachostand Null bereits in den vorausgehenden Bereinigungs-schritten eliminiert werden.

### 8.6.3 Zusammenfassung des Bereinigungs- bzw. Korrekturverfahrens

Im Zusammenhang mit der Prüfung der Plausibilität der Angaben zum Tachostand werden entsprechend dem hier entwickelten Verfahren bestimmte Fahrzeuge aus dem HU-Datenbestand eliminiert und damit von weiteren Auswertungen und Hochrechnungen ausgeschlossen. Es handelt sich dabei um Fahrzeuge,

- deren Alter zum Zeitpunkt der Prüfung Null Monate beträgt<sup>30</sup>,
- deren Tachostand 0 km oder 1 km beträgt,
- deren Tachostand unter Berücksichtigung ihres Prüfalters die untere Abschneidegrenze gemäß Tabelle 4 unterschreitet (gilt nur für Fahrzeuge, die nach dem in Tabelle 12 angegebenen Erstzulassungsjahr zugelassen worden sind),
- deren Tachostand unter Berücksichtigung ihres Prüfalters die obere Abschneidegrenze gemäß Tabelle 5 überschreitet.

Bei Fahrzeugen, die vor dem in Tabelle 12 angegeben Jahr erstmals zugelassen worden sind und deren Tachostand unter Berücksichtigung ihres Prüfalters die untere Abschneidegrenze gemäß Tabelle 10 unterschreitet, wird von einem Tachoüberlauf ausgegangen und der Tachostand wie in Abschnitt 8.6.2 beschrieben korrigiert.

## 8.7 Auswirkungen der Datenbereinigung auf das Datenvolumen

Führt man die oben beschriebenen Prüfungs-, Plausibilisierungs- und Bereinigungs-schritte konkret durch, so resultiert daraus naturgemäß eine Verringerung der Zahl der Datensätze (Fälle), die für die Auswertung und Hochrechnung zur Verfügung stehen. Insgesamt reduziert sich beispielsweise der Umfang der hier genutzten HU-Stichprobe 2007 durch das beschriebene Verfahren um 7 %. Diese Reduzierung der Fallzahl wird zu rund drei Viertel durch „ungetypte“ Kfz verursacht, die nicht in die Auswertungen einbezogen werden können.

Die Zahl der Pkw, Lkw, Sattelzugmaschinen, Omnibusse und Sonstigen Kfz in der HU-Stichprobe 2007 wird durch die Datenbereinigung nur um jeweils weniger als 2 % verringert, was für die Qualität der HU-Daten spricht. Deutlicher wirkt sich das Bereinigungsverfahren bei Kraffrädern aus, wo rund ein Zehntel der Fälle eliminiert wird; bei 92 % dieser Fälle handelt es sich im Übrigen um

---

<sup>30</sup> da mit dieser Angabe keine mittlere bisherige Monatsfahrleistung berechnet werden kann

„weitere<sup>31</sup> ungetypte“ Krafträder. Eine Analyse bei der Fahrzeugart „Zugmaschinen LoF“ – die Fahrzeugart mit der höchsten prozentualen Abnahme der Fallzahl – zeigt, dass die Verringerung der Fallzahl ganz überwiegend durch unplausibel niedrige Tachostände (1.521 von 1697 Fällen) verursacht wird. Hier kann vermutet werden, dass es sich vielfach um Angaben über Betriebsstunden statt gefahrene Kilometer handelt.

**Tabelle 14: Verringerung des Umfangs der HU-Daten 2007 durch Datenbereinigung**

Fahrzeugart	Datensätze in Gesamtrohdatendatei vor Bereinigung	Anzahl Datensätze nach letzter Bereinigung	Reduktion durch Bereinigung
Ungetypte Kfz	491017	0	- 100.0 %
Krafträder	646807	584363	-9.7 %
Pkw	7569690	7486832	-1.1 %
Lkw	486216	477003	-1.9 %
Sattelzugmaschinen	38215	37981	-0.6 %
Sonstige Zugmaschinen	16668	15441	-7.4 %
Zugmaschinen LoF	11063	9366	-15.3 %
Omnibusse	13628	13535	-0.7 %
Sonstige Kfz	29798	29365	-1.5 %
Gesamt	9303102	8653886	-7.0 %

Nach allen Bereinigungsritten liegt der Umfang der HU-Stichprobe 2007 bei 8.653.886 Kfz. Setzt man diese Zahl in Relation zu den 9.358.897 Datensätzen, die von den teilnehmenden Prüforganisationen für das Jahr 2007 ursprünglich bereitgestellt worden sind bzw. zu den 9.303.102 Datensätzen der daraus erstellten Gesamtrohdatendatei 2007, so ergibt sich daraus die bereits erwähnte Abnahme des Stichprobenumfangs insgesamt um 7 %.

Die nach dem hier entwickelten Aufbereitungsverfahren bereinigten HU-Daten 2007 und 2007 bilden die empirische Grundlage für die Entwicklung und pilotartige Erprobung eines Hochrechnungsverfahrens, welches der geplanten neuen Fahrleistungsstatistik zugrunde liegen könnte.

<sup>31</sup> Hier ist der Begriff „weitere“ so zu verstehen, dass für diese Kfz zwar die Fahrzeugart bestimmt werden kann, aber die für eine Zuordnung zu den KW-Klassen notwendigen Informationen nicht zur Verfügung stehen.



## 9 Verfahren zur Hochrechnung kalenderjahrbezogener Fahrleistungen auf der Basis von HU-Daten

### 9.1 Anforderungen an das Hochrechnungsverfahren

Unter „Hochrechnung“ ist ganz allgemein die Schätzung des unbekanntes Totalwerts  $Y$  eines Untersuchungsmerkmals auf der Basis von Stichprobendaten  $y_i$  ( $i=1, \dots, n$ ) zu verstehen. Liegt eine einfache Zufallsstichprobe vom Umfang  $n$  aus einer Grundgesamtheit vom Umfang  $N$  vor, so ist

$$\hat{Y} = (N/n) \cdot \sum y_i$$

der Schätzer für den Grundgesamtheitstotalwert  $Y$  bzw. der aus der Stichprobe „hochgerechnete“ Totalwert des betrachteten Merkmals. Der Quotient  $N/n$ , also der Kehrwert des Auswahlssatzes, wird dabei Hochrechnungsfaktor genannt.

Eine solche „freie“ Hochrechnung des kalenderjahrbezogenen Fahrleistungstotalwerts ist im Fall der HU-Daten aus mehreren Gründen nicht möglich:

- Coverage-Problem
  - Die für die Hochrechnung zur Verfügung stehenden HU-Daten eines Prüfjahres, d.h. die im betreffenden Jahr bei den datenliefernden Prüforganisationen untersuchten Fahrzeuge, sind *keine* Stichprobe aus der Grundgesamtheit aller angemeldeten Fahrzeuge. Fahrzeuge aus der Grundgesamtheit, die im betreffenden Jahr nicht zur Hauptuntersuchung vorgeführt worden sind, haben im Hinblick auf die HU-Stichprobe keine Auswahlchance.
  - Da es Grund zu der Annahme gibt, dass die kalenderjahrbezogene Fahrleistung der Fahrzeuge aus dem „nicht überdeckten“ Teil der Grundgesamtheit sich von derjenigen der im fraglichen Jahr geprüften Fahrzeuge unterscheidet, repräsentiert die vorliegende HU-Stichprobe nicht die Gesamtheit aller Fahrzeuge der Grundgesamtheit.
  - Die aus der nicht vollständigen Überdeckung der Grundgesamtheit resultierende Verzerrung der Hochrechnungsergebnisse muss durch geeignete *Korrekturfaktoren* beseitigt bzw. verringert werden.

- Fehlende Werte des eigentlichen Untersuchungsmerkmals
  - Für die Hochrechnung des kalenderjahrbezogenen Fahrleistungstotalwerts  $Y$  werden empirische Fahrleistungsdaten  $y_i$  benötigt, die sich auf das betreffende Kalenderjahr beziehen. Solche Daten liegen mit der HU-Stichprobe aber *nicht* vor, wenn man von den wenigen neuen Fahrzeugen einmal absieht, die bereits im Jahr ihrer Erstzulassung einer Hauptuntersuchung unterzogen werden.
  - Da das eigentliche Untersuchungsmerkmal bei den Einheiten in der HU-Stichprobe fehlt, weil es von den Prüforganisationen bei der Hauptuntersuchung gar nicht erhoben wird, müssen die HU-Daten vor ihrer Hochrechnung in geeigneter Weise vervollständig werden.
  - Die hierfür verwendete Missing-Data-Technik („Imputationsverfahren“) hat Einfluss auf die Genauigkeit der Hochrechnung. Im Fall der HU-Daten gilt dies in ganz besonderem Maße, da hier - als extremer Sonderfall - praktisch eine *100-Prozent-Imputation fehlender Werte* notwendig ist.
  
- Erfordernis einer nachträglichen Schichtung
  - Die periodenbezogene Fahrleistung von Fahrzeugen ist ein Merkmal mit sehr starker Streuung. In Fällen dieser Art versucht man normalerweise, die Grundgesamtheit vor der Stichprobenziehung in Teilgesamtheiten (Schichten) von Einheiten zu zerlegen, die im Hinblick auf das Untersuchungsmerkmal möglichst homogen sind.
  - Da die HU-Daten nicht aus einer geplanten Stichprobe resultieren, sondern als Prozessdaten bei den Prüforganisationen anfallen, ist die oben angedeutete a-priori-Schichtung nicht möglich. Ersatzweise kommt aber einer a-posteriori-Schichtung, d.h. eine *Zerlegung der Stichprobe in Schichten* in Betracht, die bezüglich der periodenbezogenen Fahrleistung homogen sind.
  - Als weiteres Element des Hochrechnungsverfahrens ist somit ein *Schichtungskonzept für die HU-Stichprobe* zu entwickeln. Die betreffenden Schichtungsvariablen zur Abgrenzung von „Hochrechnungsgruppen“ müssen Variable aus dem HU-Datensatz sein, die mit der Zielvariablen „periodenbezogene Fahrleistung“ bzw. dem dafür verwendeten Surrogat korreliert sind. Zugleich müssen die betreffenden Variablen aber auch im ZFZR (als dem Verzeichnis der zur Grundgesamtheit gehörenden Einheiten) vorhanden sein, damit ein entsprechender „Hochrechnungsrahmen“ gebildet werden kann.

Ein statistischer Hochrechnungsansatz, der den genannten Besonderheiten der HU-Daten Rechnung trägt, wird im Folgenden hergeleitet. Die Darstellungen beziehen sich auf das *mathematisch-statistische Verfahren*, nach welchem die Hochrechnung erfolgt. Davon zu unterscheiden ist die Beschreibung der *rechentechnischen Umsetzung* des Verfahrens mit Hilfe von SPSS; diese liegt in Band 2 des Schlussberichts in Form einer Übersichtsdarstellung (Kapitel 1) und im programmier-technischen Detail als SPSS-Syntax (Kapitel 3) vor.

## 9.2 Prinzip der Übereinstimmung von HU-Prüfjahr und Bezugsjahr der Hochrechnung

Das Hochrechnungsverfahren beruht auf dem Grundsatz, dass zur Schätzung des Fahrleistungstotalts für ein bestimmtes Kalenderjahr die HU-Prüfdaten des betreffenden Kalenderjahres als empirische Basis dienen. Da die Prüftermine der Fahrzeuge über das gesamte Kalenderjahr verteilt sind, variiert die Länge der im betrachteten Kalenderjahr liegenden „aktuellen“ Periode der einzelnen Fahrzeuge in der HU-Stichprobe zwischen 0 Monaten (z.B. Prüftermin 3. Januar) und 12 Monaten (z.B. Prüftermin 22. Dezember).

Natürlich könnte man auch daran denken, für die Schätzung des auf das Kalenderjahr  $t$  bezogenen Fahrleistungstotalts zusätzlich auch HU-Daten aus dem nachfolgenden Kalenderjahr  $t+1$  (Prüftermin liegt im Kalenderjahr  $t+1$ ) zu verwenden. Würde man unter dem „aktuellen Zeitraum“ dann z.B. die letzten 6 Monate vor dem Prüftermin verstehen, so würde für die in der ersten Hälfte des Jahres  $t+1$  geprüften Fahrzeuge der individuelle aktuelle Zeitraum in das Kalenderjahr  $t$  (dem Bezugsjahr der Hochrechnung) zurückreichen.

Da es bei diesem Ansatz notwendig wäre, mit der Schätzung für das Bezugsjahr  $t$  so lange zu warten, bis die HU-Daten aus dem darauf folgenden Jahr  $t+1$  vorliegen, wurde unter dem Gesichtspunkt eine möglichst hohen Aktualität der Fahrleistungsschätzung ein solches Hochrechnungskonzept nicht weiter verfolgt.

## 9.3 Grundstruktur des Hochrechnungsverfahrens

### 9.3.1 Hochrechnungsrelevante Variablen

Das Ausgangsdatenmaterial der Hochrechnung des Totalwerts der Fahrleistung im Kalenderjahres  $t$  ist die *Datenmatrix*  $D_0$  mit den Einzeldaten der aufbereiteten HU-Stichprobe des Prüfjahres  $t$ . Jede Zeile  $i$  ( $i=1, \dots, n$ ) der Datenmatrix repräsentiert ein geprüftes Fahrzeug, jede Spalte ein Merkmal des Fahrzeugs  $i$ . Die in der geprüften und aufbereiteten HU-Rohdatendatei enthaltenen Merkmale sind bereits in Kapitel 8 dargestellt worden. Eine detaillierte Liste („list of variables on the working file“) findet sich in Band 2 des Schlussberichts (Kapitel 2).

Mit Blick auf die Hochrechnung stehen folgende Merkmale des Fahrzeugs  $i$  ( $i=1, \dots, n$ ) im Mittelpunkt des Interesses (Spalten der Datenmatrix):

- Erstzulassungsjahr des Fahrzeugs ( $e_i$ )
- Alter des Fahrzeugs (in Jahren) im Bezugsjahr der Hochrechnung ( $a_i = t - e_i$ )
- Alter des Fahrzeugs (in Monaten) beim Prüftermin ( $k_i$ )
- Tachostand (in km) beim Prüftermin ( $z_i$ )
- Index der Hochrechnungsschicht zu welcher das Fahrzeug  $i$  gehört ( $h_i$ )
- Fahrzeugmerkmale nach denen die hochgerechnete Fahrleistung aufgegliedert<sup>32</sup> werden soll ( $v_{1i}, v_{2i}, v_{3i}, \dots$ )
- mittlere Fahrleistung (in km/Monat) seit Erstzulassung<sup>33</sup> ( $s_i = z_i / k_i$ ).

Nach dem in Abschnitt 9.1 beschriebenen methodischen Grundkonzept muss die Datenmatrix  $D_0$  um weitere Variable ergänzt werden, damit die Hochrechnung konkret durchgeführt werden kann. Im Kern handelt es sich dabei um die folgenden Variablen:

- eine aus den Variablen  $s_i$ ,  $h_i$  und  $e_i$  sowie gewissen exogenen Daten (FU-Statistik und Kfz-Bestandsstatistik) gebildete *modifizierte Fahrleistungsvariable*  $\hat{s}_i$  (sog. „neutrale“ oder „androgyn“ mittlere Fahrleistung in km/Monat), die anstelle der ursprünglichen Fahrleistungsvariablen  $s_i$  als „Zielvariable“ bei der Hochrechnung verwendet wird, um die aus der Nichterfassung ungeprüfter Fahrzeuge in der HU-Stichprobe resultierende Verzerrung der Schätzung des Fahrleistungstotals („Coverage-Fehler“) zu korrigieren (vgl. Abschnitt 9.5.4) sowie
- ein *Hochrechnungsfaktor* ( $g_i$ ) in der Einheit „Gesamtzahl Fahrzeugmonate“<sup>34</sup> in der Grundgesamtheit pro Fahrzeug in der HU-Stichprobe“, der von der Schichtzugehörigkeit ( $h_i$ ) sowie dem Erstzulassungsjahr ( $e_i$ ) des Fahrzeugs abhängt und zu dessen Berechnung exogene Daten aus der Kfz-Bestandsstatistik benötigt werden.

---

<sup>32</sup> Als Aufgliederungsmerkmale kommen im Prinzip alle im Rohdatensatz enthaltenen Merkmale einschließlich der zur Abgrenzung der Hochrechnungsschichten verwendeten Merkmale infrage (z.B. Fahrzeugart, Hersteller, Antriebsart etc.).

<sup>33</sup> Nach der Durchschnittswertmethode erzeugte („imputierte“) künstliche Zielvariable, die an die Stelle der fehlenden Variablen „Fahrleistung je Monat im Bezugsjahr der Hochrechnung“ tritt.

<sup>34</sup> Fahrzeugmonate im Bezugsjahr der Hochrechnung

Die Methode zur Berechnung der modifizierten Fahrleistungsvariablen  $\hat{s}_i$  und des Hochrechnungsfaktors  $g_i$  wird in den nachfolgenden Abschnitten dargestellt. Die um diese beiden neuen Variablen (Spalten) erweiterte Datenmatrix  $D_1$  ist die empirische Grundlage der Hochrechnung.

### 9.3.2 Schätzfunktion für den kalenderjahrbezogenen Fahrleistungstotalwert

Liegt die erweiterte Datenmatrix  $D_1$  vor, so kann der Schätzwert für das Fahrleistungstotal im betrachteten Kalenderjahr auf einfache Weise berechnet werden. Als Schätzfunktion  $\hat{Y}$  für das unbekannte Fahrleistungstotal  $Y$  im betrachteten Kalenderjahr dient die gewichtete Summe der modifizierten Fahrleistungsvariablen, d.h. die Kennzahl

$$(9-1) \quad \hat{Y} = \sum \hat{s}_i \cdot g_i$$

wobei über alle Fahrzeuge  $i=1, \dots, n$  in der HU-Stichprobe (Zeilen der erweiterten Datenmatrix  $D_1$ ) summiert wird. Das geschätzte Fahrleistungstotal  $\hat{Y}$  wird dabei in der Einheit km/Jahr gemessen, d.h.

$$[\hat{Y}] = \text{km/Jahr.}$$

Diese Einheit ergibt sich aus dem Produkt von

$$[\hat{s}_i] = \text{km/Monat} \quad \text{und} \quad [g_i] = (\text{Fahrzeug-Monat /Jahr})/\text{Fahrzeug.}$$

Bei Ergebnisdarstellungen ist es naheliegend, das geschätzte Fahrleistungstotal eines bestimmten Kalenderjahres in der Einheit Mrd. km/Jahr oder Mio. km/Jahr anzugeben, um große Zahlenwerte in den entsprechenden Tabellen zu vermeiden. Hinzu kommt, dass die Einheit km/Jahr eine Ergebnissenauigkeit suggeriert, die nach Lage der Dinge nicht sachgerecht ist.

### 9.3.3 Fahrleistungsschätzung für Teilgruppen von Fahrzeugen

Soll das kalenderjahrbezogene Fahrleistungstotal nicht nur für die Gesamtheit aller Fahrzeuge sondern auch für Teilgruppen von Fahrzeugen geschätzt werden, so ist die Hochrechnungsformel (9-1) getrennt für jede Teilgruppe anzuwenden.

Gliedert man die  $n$  Fahrzeuge in der Stichprobe z.B. nach  $K$  Fahrzeugarten und  $L$  Antriebsarten auf, so entstehen  $K \cdot L$  Teilgruppen. Das unbekannte Fahrleistungstotal  $Y(j)$  der Teilgruppe  $G(j)$  ( $j=1, \dots, K \cdot L$ ) wird dann durch

$$(9-2) \quad \hat{Y}(j) = \sum_{i \in G(j)} \hat{s}_i \cdot g_i$$

geschätzt (Summation über alle Fahrzeuge der Teilgruppe  $G(j)$ ).

Aufgliederungen des Fahrleistungstotalts sind mit dem beschriebenen Verfahren nicht nur nach Schichtungsmerkmalen (Merkmale zur Abgrenzung der Hochrechnungsgruppen) möglich, sondern prinzipiell nach jedem Merkmal, das als Variable (Spalte) in der Datenmatrix enthalten ist. Hier sind speziell auch die Variablen angesprochen, welche über die Hersteller-Typ-Referenzdatei dem HU-Datensatz nachträglich hinzugefügt werden können (vgl. Abschnitt 8.4).

## 9.4 Hochrechnungsgruppen: Nachträgliche Schichtung der HU-Stichprobe

Zur Verbesserung der Hochrechnungsgenauigkeit wird die HU-Stichprobe nach dem Konzept der nachträglichen Schichtung (auch „Redressment“ genannt) in sog. Hochrechnungsgruppen zerlegt. Grundlage hierfür ist ein *hierarchisches* Schichtungskonzept mit dem Merkmal „Fahrzeugart“ auf der obersten Ebene und weiteren Schichtungsmerkmalen auf nachgelagerten Ebenen. Das Merkmal „Fahrzeugalter im Bezugsjahr“ dient dabei in allen Fahrzeugarten als weiteres Schichtungsmerkmal (wenngleich unterschiedlich skaliert). Das Konzept trägt den generellen statistischen Anforderungen an Schichtungsmerkmale Rechnung und berücksichtigt zugleich, dass die einzelnen Fahrzeuggruppen in der Stichprobe (und ebenso in der Grundgesamtheit) sehr unterschiedlich stark besetzt sind.

### 9.4.1 Schichtungsmerkmal „Fahrzeugart“

Die Fahrzeugart rangiert als Schichtungsmerkmal aus zwei verschiedenen Gründen auf der obersten Ebene:

- Durch die Unterscheidung nach Fahrzeugarten entstehen „homogene“ Hochrechnungsgruppen, da es zwischen den Fahrzeugarten beträchtliche Mittelwertunterschiede im Hinblick auf das Merkmal Fahrleistung gibt.
- Bei nahezu allen Ergebnistabellen einer Fahrleistungsstatistik ist aus inhaltlich-sachlogischen Gründen nach Fahrzeugarten zu differenzieren. Durch die nachträgliche Schichtung der HU-Stichprobe nach der Fahrzeugart kann sichergestellt werden, dass die im Rahmen der Hochrechnung verwendeten Kfz-Bestandszahlen mit der amtlichen Kfz-Bestandsstatistik übereinstimmen.

Zur konkreten Erstellung des Hochrechnungsrahmens muss durch das Kraffahrt-Bundesamt der ZFZR-Bestand je Fahrzeugschicht (vgl. Abschnitt 9.6) ausgezählt werden. Ab dem Jahr 2008 enthält das Register nur noch die Fahrzeuge des „aktiven“ Bestands (d.h. ohne vorübergehend stillgelegte Kfz). Im Rahmen der pilotartigen Anwendung und Erprobung des Hochrechnungsverfahrens erfolgte eine Auszählung des Kfz-Bestands für die Stichtage 1.1.2006, 1.1.2007 und 1.1.2008, wobei unterstellt wird, dass der Bestand zum 1.1. eines Jahres jeweils mit dem Bestand zum 31.12. des Vorjahres gleichzusetzen ist. Die Ergebnisse dieser Sonderauszählung für den 1.1.2008 sind - zusammengefasst nach Fahrzeugarten - den von der amtlichen Statistik veröffentlichten Zahlen (Statistische Mitteilungen des Kraffahrt-Bundesamtes, FZ 6, 1. Januar 2008) in der folgenden Tabelle gegenübergestellt. Meist stimmen die Werte genau oder nahezu überein, die Unterschiede bei „Sonstige Zugmaschinen“ und „Zugmaschinen LoF“ können bei zukünftigen Auswertungen ggf. vermieden werden.

**Tabelle 15: Vergleich der zur Erstellung eines Hochrechnungsrahmens für die HU-Stichprobe durchgeführten Sonderauszählung des Kfz-Bestands mit veröffentlichten Kfz-Bestandszahlen (1.1.2008)**

Fahrzeugart	Sonderauszählung (A)	Veröffentl. Kfz-Bestand (B)	Relative Abweichung $[(A)-(B)]/(A)$	Abweichung (A) – (B)
Krafträder	3.551.724	3.566.122	-0,4 %	-14.398
Pkw	41.179.747	41.183.594	0,0 %	-3.847
Lkw	2.323.052	2.323.064	0,0 %	-12
Sattelzugmaschinen	179.935	179.935	0,0 %	0
Sonstige Zugmaschinen	691.489	588.427	14,9 %	103.062
Zugmaschinen LoF	1.051.791	1.154.873	-9,8 %	-103.082
Omnibusse	75.068	75.068	0,0 %	0
Sonstige Kfz	262.094	258.954	1,2 %	3.140
Gesamt	49.314.900	49.330.037	0,0 %	-15.137

#### 9.4.2 Schichtungsmerkmal „Zulassungsjahr“ bzw. „Fahrzeualter“

Bei der empirischen Grundlage der Fahrleistungsschätzung, den HU-Daten eines bestimmten Prüffjahres, handelt es sich um Querschnittsdaten, wenn auch der Zeitraum, in welchem die betreffenden Hauptuntersuchungen stattgefunden haben, jeweils ein ganzes Kalenderjahr umfasst. Aus anderen Querschnittsuntersuchungen (z.B. FLE 2002, KiD 2002) ist bekannt, dass es zwischen der aktuellen Fahrleistung von Fahrzeugen (in einem kürzeren Zeitraum vor dem Analysezeitpunkt) und dem Alter der Fahrzeuge (Alter zum Zeitpunkt der Analyse) einen statistischen Zusammenhang gibt: Untersucht man zu einem gegebenen Analysezeitpunkt ein bestimmtes Fahrzeugkollektiv, so stellt man eine negative Korrelation zwischen der aktuellen Fahrleistung und dem zum Analysezeitpunkt erreichten Fahrzeualter fest. Da es in einer laufenden Fahrleistungsstatistik

auf der Basis von HU-Daten um die Schätzung der Fahrleistung im aktuellen Kalenderjahr geht, liegt es nahe, bei der Abgrenzung von Hochrechnungsgruppen das Merkmal „Alter des Fahrzeugs“ als Schichtungsmerkmal zu verwenden.

Dabei gibt es nun mehrere Möglichkeiten, die Länge des Vergangenheitszeitraums, in welchem bereits Fahrleistungen erbracht worden sind, operational zu beschreiben. Bei der Entscheidung für eine bestimmte Operationalisierung ist natürlich auch die Möglichkeit zur Verknüpfung der HU-Daten mit ZFZR-Bestandsdaten zu berücksichtigen, da der Hochrechnungsrahmen von ZFZR-Daten zum jahresdurchschnittlichen Bestand angemeldeter Fahrzeuge gebildet wird.

Im Hinblick auf die Bildung von Hochrechnungsgruppen ist zur Charakterisierung der bisherigen Lebensdauer von Fahrzeugen dem Merkmal „(Erst-)Zulassungsjahr“<sup>35</sup> gegenüber dem Merkmal „Prüfalter“<sup>36</sup> der Vorzug zu geben. Wie man sich nämlich leicht überlegt, kann ein im Jahr 2007 geprüftes Fahrzeug, das ein „Prüfalter“ von z.B. 6 Monaten aufweist, sowohl aus der Kohorte 2007 (Untersuchungsjahr = Zulassungsjahr) als auch aus der Kohorte 2006 (Untersuchungsjahr = Zulassungsjahr +1) stammen. Demgegenüber ist das Zulassungsjahr ein zeitinvariantes Fahrzeugmerkmal im HU-Datenbestand, welches direkt mit dem entsprechenden ZFZR-Merkmal korrespondiert. Das Zulassungsjahr ist als Schichtungsmerkmal geeignet, weil die Zielgröße der Hochrechnung (nämlich die Fahrleistung in der aktuellen Periode, ersatzweise berechnet als mittlere bisherige Fahrleistung) mit dem Merkmal Zulassungsjahr korreliert.

Anmerkung: Für Zwecke der Darstellung des Schichtungsplans und des Hochrechnungsverfahrens insgesamt - aber auch im Hinblick auf die Darstellung von Hochrechnungsergebnissen - ist es vielfach anschaulicher, nicht vom „Zulassungsjahr“ eines Fahrzeugs sondern vom „Alter des Fahrzeugs im Bezugsjahr der Hochrechnung (in Jahren)“ zu sprechen.

Unter dem Alter des Fahrzeugs ist dann ganz einfach die Differenz  $a=t-e$  zu verstehen, wobei  $t$  das Bezugsjahr der Hochrechnung und  $e$  das Erstzulassungsjahr des Fahrzeugs bezeichnet. Die beiden Merkmale entsprechen also einander.

Die folgende Abbildung 7, die basierend auf den HU-Daten des Jahres 2007 für fünf verschiedene Fahrzeugarten die mittlere bisherige Monatsfahrleistung in Abhängigkeit vom Zulassungsjahr des Fahrzeugs zeigt, lässt die negative Korrelation zwischen Schichtungsmerkmal und Zielgröße der Hochrechnung erkennen.

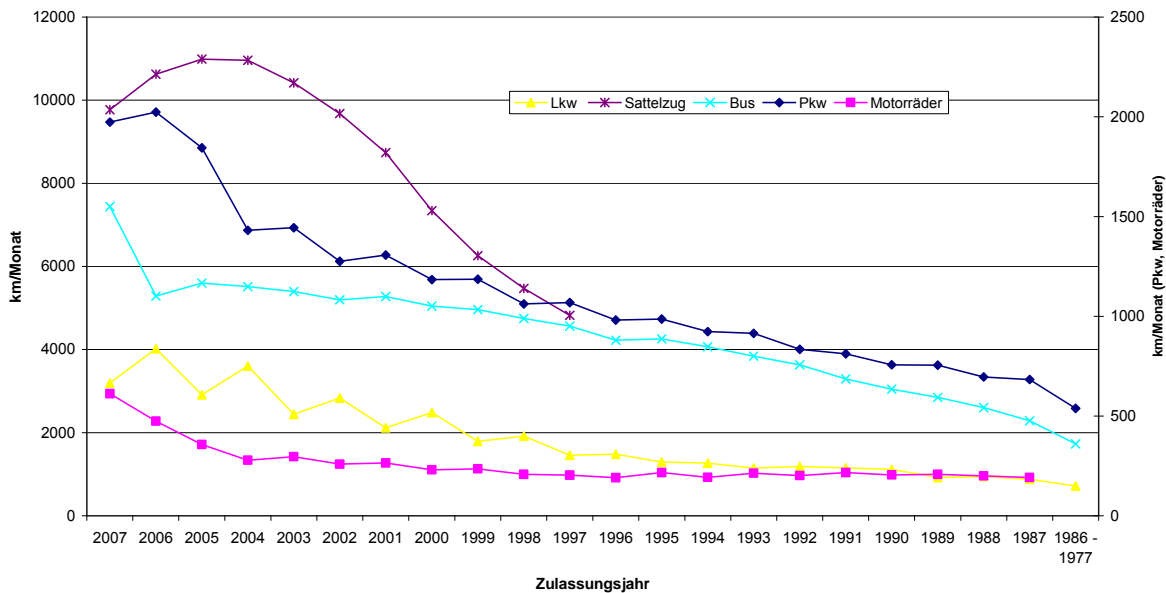
---

<sup>35</sup> Im vorliegenden HU-Datenbestand hat dieses Merkmal die Ausprägungen 2007, 2006, 2005, ...

<sup>36</sup> Alter des Fahrzeugs zum Zeitpunkt der betreffenden Hauptuntersuchung



**Abbildung 7: Mittlere bisherige Monatsfahrleistung von im Jahr 2007 geprüften Fahrzeugen nach Fahrzeugart und Zulassungsjahr**



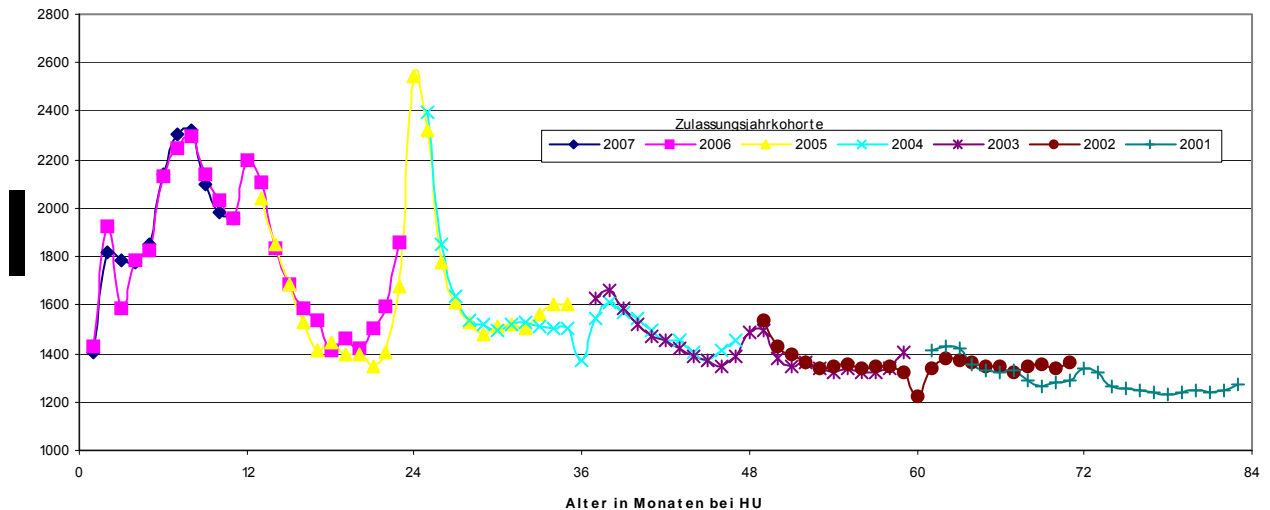
### Ergänzende Anmerkungen

Ergänzend zeigt die folgende Abbildung 8 für das HU-Prüfjahr 2007 und die Fahrzeugart „Pkw“ die mittlere bisherige Monatsfahrleistung in Abhängigkeit vom Alter des Fahrzeugs (in Monaten) bei der Prüfung, wobei zusätzlich nach dem Zulassungsjahr (Kohorten 2007 – 2001) unterschieden wird. Wie man sieht, stimmen bei gegebenem Alter zum Prüfzeitpunkt im Jahr 2007 die jeweiligen kohortenspezifischen Durchschnittsfahrleistungen sehr gut überein.

Eine kurze Erläuterung soll die Interpretation der Abbildung erleichtern. Betrachtet man beispielsweise die Kohorten 2006 und 2005, so entsprechen sich die mittleren bisherigen Monatsfahrleistungen bei gleichem Prüfalter (im Jahr 2007) ziemlich genau. Der Überlappungsbereich des Fahrzeugalters in Monaten beträgt bei diesen beiden Kohorten 13 bis 23 Monate (Alter des Fahrzeugs zum Zeitpunkt der Prüfung im Jahr 2007).

Es muss aber darauf hingewiesen werden, dass bei anderen Fahrzeugarten die Übereinstimmung der Kohorten nicht immer so deutlich ist (z.B. bei Krafträdern). Wenn bei gleichem Alter zum Prüfzeitpunkt (der im Jahr 2007 liegt) sich zwischen den beiden „überlappenden“ Kohorten die mittlere bisherige Monatsfahrleistung unterscheidet, dann dürfte dies weniger an zulassungsjahrspezifischen Besonderheiten liegen, sondern vielmehr daran, dass selbst bei einem Stichprobenumfang pro Fahrzeuggruppe von insgesamt 500.000 Fahrzeugen einige Altersmonate in den einzelnen Kohorten nur sehr schwach besetzt sind.

**Abbildung 8: Mittlere bisherige Monatsfahrleistung von Pkw gegliedert nach Zulassungsjahr und Alter bei der Hauptuntersuchung in Monaten (HU-Prüfjahr 2007)**



### 9.4.3 Weitere Merkmale für die Bildung von Hochrechnungsgruppen

Neben „Fahrzeugart“ und „Fahrzeugalter im Bezugsjahr der Hochrechnung“ (abgeleitet aus dem Jahr der Zulassung) wurden für die Fahrzeugarten Kraffräder, Pkw und Lkw weitere Schichtungsmerkmale für die Bildung von Hochrechnungsgruppen festgelegt. Für jede der drei genannten Fahrzeugarten wurde als weiteres Schichtungsmerkmal ein geeignetes Leistungs- bzw. Größenmerkmal festgelegt, da insbesondere aus der Fahrleistungserhebung 2002 bekannt ist, dass die Fahrleistung von Fahrzeugen tendenziell mit der Motorleistung und mit der Größe (Ladefähigkeit) steigt. Bei der Fahrzeugart Pkw wird als viertes Schichtungsmerkmal zusätzlich noch die Antriebsart eingeführt. Zum einen gibt es ein generelles Interesse an Pkw-Fahrleistungen nach Antriebsart (ganz aktuell etwa im Hinblick auf Hybridfahrzeuge). Zum anderen ist die Unterscheidung nach der Antriebs- und damit der Kraftstoffart (Benzin, Diesel) für die Zusammenführung der Hochrechnung der HU-Stichprobe mit den Resultaten der bereits erwähnten DIW-Fahrleistungsermittlung nach dem Verfahren der „Kraftstoffverbrauchsrückrechnung“ von Bedeutung.

### 9.4.4 Der Schichtungsplan im Überblick

Der endgültige Schichtungsplan für die Hochrechnung der HU-Stichprobe entsteht durch Kombination der verschiedenen Schichtungsmerkmale. Jede einzelne Merkmalskombination entspricht dabei einer Hochrechnungsgruppe. Beispiele für Hochrechnungsgruppen sind etwa die Merkmalskombinationen „Pkw / 2 Jahre / Diesel / 45 bis 59 KW“ und „Lkw / 21 bis 30 Jahre / über 7,5 t zGG“. In der nachfolgenden Tabelle 16 sind die Schichtungsmerkmale und ihre Ausprägungen im

Überblick dargestellt. Bei der Spezifizierung der Ausprägungen wurde, wie erwähnt, auch darauf geachtet, dass die resultierenden gruppenspezifischen Stichprobenumfänge nicht zu klein werden.

**Tabelle 16: Schichtungsplan zur Abgrenzung der Hochrechnungsgruppen**

Fahrzeugart	Fahrzeugalter im Bezugsjahr der Hochrechnung (in Jahren)	Antriebsart	Leistungs- bzw. Größenmerkmal
Krafträder	0, 1, 2, ..., 30, 31 bis 50, 51+	-	bis 15 KW , 16 bis 64 KW, 65+ KW
Pkw	0, 1, 2, ..., 20, 21 bis 30, 31+	Diesel	bis 44 KW, 45 bis 59 KW, 60 bis 89 KW , 90+ KW
Pkw	0, 1, 2, ..., 20, 21 bis 30, 31+	Benzin	bis 44 KW, 45 bis 59 KW, 60 bis 89 KW , 90+ KW
Pkw	0, 1, 2, ..., 20, 21 bis 30, 31+	Hybrid/sonstiger Antrieb	-
Lkw	0, 1, 2, ..., 20, 21 bis 30, 31+	-	bis 3,5 t zGG, über 3,5 bis 7,5 t zGG , über 7,5 t zGG
Sattelzugmaschinen	0, 1, 2, ..., 10, 11 bis 20, 21+	-	-
Sonstige Zugmaschinen	0, 1, 2, ..., 30, 31 bis 50, 51+	-	-
Zugmaschinen LoF	0, 1, 2, ..., 30, 31 bis 50, 51+	-	-
Omnibusse	0, 1, 2, ..., 20, 21 bis 30, 31+	-	-
Sonstige Kfz	0, 1, 2, ..., 25, 26 bis 40, 41+	-	-

## 9.5 Korrektur des Coverage-Fehlers: Konzept der „neutralen“ Fahrleistung

### 9.5.1 Ursachen und Folgen des Coverage-Fehlers

Geht man von der Vorstellung aus, dass die Grundgesamtheit der Fahrzeuge eines bestimmten Jahres in die beiden Teilgesamtheiten der im betreffenden Jahr geprüften bzw. nicht geprüften Fahrzeuge zerfällt, so kann man sagen, dass die HU-Stichprobe eines Prüffjahres einen beträchtlichen Teil der Grundgesamtheit aller Fahrzeuge nicht überdeckt. Sofern ein Fahrzeug zur Teilgesamtheit der nicht geprüften Fahrzeuge gehört, hat es keine Chance, in der HU-Stichprobe zu erscheinen.

Zwischen den in einem bestimmten Jahr geprüften und den nicht geprüften Fahrzeugen gibt es im Hinblick auf einige hochrechnungsrelevante Merkmale beträchtliche Strukturunterschiede. Diese betreffen auf jeden Fall das Zulassungsjahr (bzw. Fahrzeugalter im Bezugsjahr), aber auch die gesamte bisherige und die periodenbezogene Fahrleistung. Ihre Ursache haben diese Strukturunterschiede einerseits in den Vorschriften zu den Intervallen, in denen Fahrzeuge zur Hauptuntersuchung vorgeführt werden müssen und andererseits im Kfz-Nutzungsverhalten sowie den weiteren Nutzungsabsichten der Fahrzeughalter.

Gemäß Anlage VIII zu § 29 StVZO sind die Untersuchungsintervalle für die Fahrzeugarten unterschiedlich geregelt:

- Für Pkw gilt bei Erstzulassung ein Untersuchungsintervall von 36 Monaten, ab der ersten HU ein Intervall von 24 Monaten.
- Für Motorräder und Lkw bis 3,5 t zulässigem Gesamtgewicht gilt ein Untersuchungsintervall von generell 24 Monaten,
- für alle anderen Nutzfahrzeuge (inkl. Omnibusse) gilt ein Untersuchungsintervall von 12 Monaten.

Ausnahmen von diesen Regelungen gibt es bei der Fahrzeugart Pkw für Taxis und Mietwagen; deren Untersuchungsintervall beträgt generell 12 Monate. Eine weitere Ausnahme bilden Fahrzeuge zur Güterbeförderung/Zugmaschinen, deren Höchstgeschwindigkeit nicht mehr als 40 km/h beträgt; die Untersuchungsintervalle betragen hier generell 24 Monate. Schließlich werden Wohnmobile bis 3,5 t zulässigem Gesamtgewicht wie Pkw geprüft, bei Wohnmobilen zwischen 3,5 t und 7,5 t zulässigem Gesamtgewicht betragen die Untersuchungsintervalle immer 24 Monate und bei Wohnmobilen ab 7,5 t zulässigem Gesamtgewicht immer 12 Monate.

Als Folge der genannten Vorschriften ist es mit Blick auf das HU-Prüfjahr 2007 beispielsweise so, dass Pkw der Zulassungsjahre (Kohorten) 2007, 2006 und 2005 gegenüber Pkw des Zulassungsjahrs 2004 in der HU-Stichprobe 2007 deutlich seltener vertreten und damit gemessen an der Grundgesamtheit aller Pkw weit unterrepräsentiert<sup>37</sup> sind. Die reine Unter- bzw. Überrepräsentierung einer Pkw-Kohorte erscheint zunächst nicht besonders problematisch, da ja mit Hilfe der nach Zulassungsjahren gegliederten Pkw-Bestandszahlen 2007 eine entsprechende Anpassung (Gewichtung) der HU-Stichprobe möglich ist.

Zum Problem wird der geschilderte Sachverhalt dadurch, dass die in der HU-Stichprobe eines bestimmten Jahres enthaltenen (geprüften) Fahrzeuge einer bestimmten Kohorte im Hinblick auf die

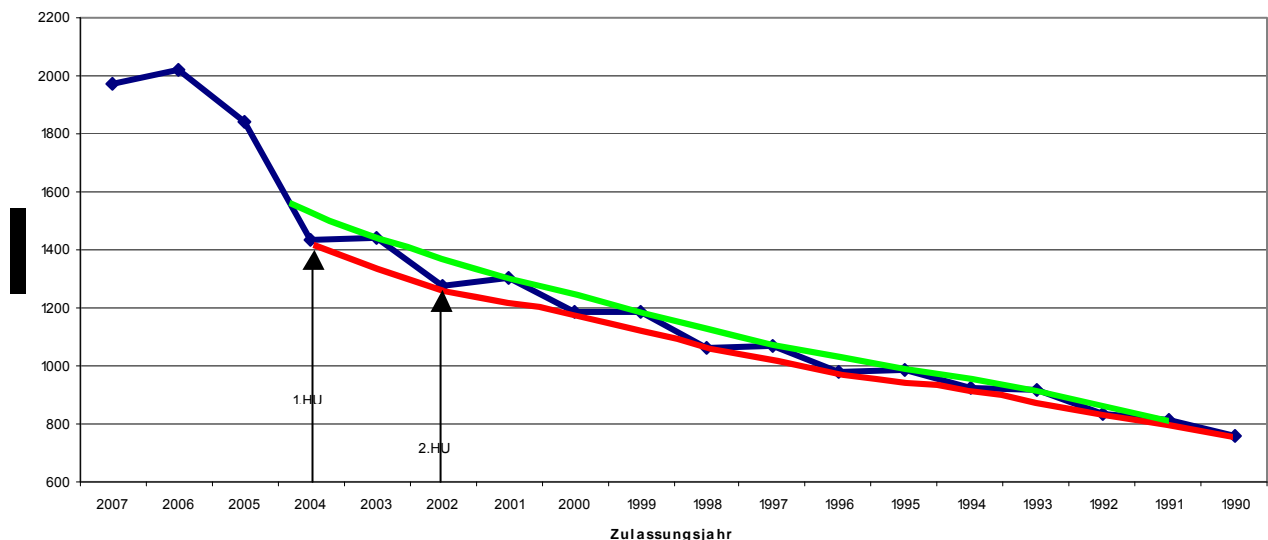
---

<sup>37</sup> Spiegelbildlich sind Pkw des Zulassungsjahrs 2004 in der HU-Stichprobe 2007 stark überrepräsentiert.

Fahrleistung *nicht repräsentativ* sind für alle zur Grundgesamtheit gehörenden Fahrzeuge der betreffenden Kohorte. Es ist beispielsweise so, dass der im Jahr 2007 zu einer Hauptuntersuchung vorgeführte Teil der Pkw-Kohorte 2006 schon allein deshalb überdurchschnittlich hohe Fahrleistungen aufweist, weil wegen der kürzeren Prüfungsintervalle für Taxen und Mietwagen diese unter den 2007 geprüften Pkw der Kohorte 2006 weit überproportional vertreten sind.

Die nachfolgende Abbildung 9 zeigt für das HU-Prüfjahr 2007 und die Fahrzeugart Pkw die durchschnittliche bisherige Monatsfahrleistung in Abhängigkeit vom Zulassungsjahr. Es zeigt sich deutlich, dass ab dem ersten (regelmäßigen) Intervall von 36 Monaten im Prinzip zwei unterschiedliche „Fahrleistungsniveaus“ vorliegen, wobei die Pkw mit den „regelmäßigen“ Untersuchungsintervallen (hier vertreten durch die Kohorten 2004, 2002, 2000 etc.) eine niedrigere durchschnittliche Monatsfahrleistung aufweisen als die Pkw mit „unregelmäßigen“ Untersuchungsintervallen (hier vertreten durch die Kohorten 2007, 2006, 2005, 2003, 2001 etc.). Für Pkw, die vor dem ersten regelmäßigen Prüftermin zur Hauptuntersuchung vorgeführt werden (Kohorten 2007, 2006 und 2005) zeigen sich wegen der überproportional häufig vertretenen Taxen, Miet- und Geschäftswagen besonders stark nach oben abweichende Fahrleistungen.

**Abbildung 9: Mittlere bisherige Monatsfahrleistungen von regelmäßig und unregelmäßig geprüften Pkw nach Zulassungsjahr (HU-Prüfdaten 2007)**



## 9.5.2 Folgerungen für das Hochrechnungsverfahren

Wie gezeigt wurde, gibt es bei gegebenem Prüf- und Erstzulassungsjahr einen statistischen Zusammenhang zwischen dem Merkmal „Prüfstatus“, welches darüber entscheidet, ob ein Fahrzeug eine von Null verschiedene Auswahlchance für die HU-Stichprobe hat oder nicht, und dem Merk-

mal „bisherige Fahrleistung“, welches die Zielvariable der Hochrechnung darstellt. Zur Vermeidung bzw. Verringerung der aus dem oben geschilderten Zusammenhang resultierenden Verzerrungen der Fahrleistungsschätzung sind an das Hochrechnungsverfahren folgenden Anforderungen zu stellen:

1. Fahrzeuge, die vor dem ersten regelmäßigen Prüftermin einer Hauptuntersuchung unterzogen werden, sind bei der Hochrechnung gesondert zu betrachten. Greift man als Beispiel das Prüf- bzw. Hochrechnungsjahr 2007 heraus, so gilt dies bei „Sattelzugmaschinen“, „Lkw ab 3,5 t zulässigem Gesamtgewicht“ und „Omnibussen“ für die Kohorte 2007, bei „Krafträder“ und „Lkw bis 3.5 t zulässigem Gesamtgewicht“ für die Kohorten 2007 und 2006 und bei Pkw für die Kohorten 2007, 2006 und 2005.
2. Bei Fahrzeugarten, deren regelmäßige Untersuchungsintervalle länger als 12 Monate sind, sollten - sofern „alternierende“ mittlere Fahrleistungen (aufeinanderfolgender Kohorten) festzustellen sind - entsprechende Verzerrungskorrekturen vorgenommen werden. Aufgrund von Analysen der HU-Stichprobendaten (nach dem Schema gemäß Abbildung 10 betrifft dies die Fahrzeugarten Pkw, Krafträder und Lkw bis 3,5 t zulässigem Gesamtgewicht. Eine Verzerrungskorrektur ist möglich, wenn für das Bezugsjahr der Hochrechnung der Grundgesamtheitsanteil der geprüften Kfz je Kohorte (Zulassungsjahr) bekannt ist.

Das zur Korrektur von Verzerrungen entwickelte Konzept der „neutralen“ Fahrleistung eines (geprüften) Fahrzeugs aus der HU-Stichprobe wird nachfolgend dargestellt.

### 9.5.3 Grundgedanke des Korrekturverfahrens

Der Grundgedanke des Verfahrens zur Korrektur des Coverage-Fehlers soll am Beispiel der Fahrzeugart Pkw unter Bezugnahme auf Abbildung 9 verdeutlicht werden.

Betrachtet man die Pkw-Kohorte 2003, so handelt es sich um eine Kohorte, die im Prüffahr 2007 nicht zu den „regelmäßig“ geprüften Fahrzeugen gehört, da die betreffenden Pkw in diesem Jahr im Durchschnitt 48 Monate alt sind. Mit Hilfe der FU-Statistik 2007, die ja eine Totalerhebung der in Deutschland 2007 durchgeführten Hauptuntersuchungen von Pkw darstellt (Gesamtzahl rund 18 Mio. gemäß Tabelle 17) und der ZFZR-Daten zum Fahrzeugbestand lässt sich für die Pkw-Kohorte 2003 und das Prüffahr 2007 ein Prüfanteil von 26 % des jahresdurchschnittlichen Bestands dieser Kohorte ermitteln<sup>38</sup>. Für diese Pkw - die aus der Kohorte 2003 stammen und 2007 einer Hauptuntersuchung unterzogen worden sind - sind aus der HU-Stichprobe 2007 die mittleren bisherigen

---

<sup>38</sup> Eine allgemeine Darstellung hierzu findet sich in Abschnitt 9.5.5.

Monatsfahrleistungen bekannt; für die komplementären 74 % der im Jahr 2007 *nicht* geprüften Fahrzeuge der Kohorte 2003 können dagegen zunächst keine Aussagen hinsichtlich der Fahrleistung getroffen werden. Über diese Fahrzeuge müssen im Rahmen des Hochrechnungsverfahrens also gewisse Annahmen getroffen werden.

Hier liegt die Annahme nahe, dass die mittlere bisherige Fahrleistung der im Jahr 2007 *nicht geprüften* 74 % der Pkw-Kohorte 2003<sup>39</sup> näherungsweise gleich dem Durchschnitt der Fahrleistungen der ein Jahr jüngeren (2004) und der ein Jahr älteren (2002) Kohorte ist. Mit den Daten dieser beiden Kohorten<sup>40</sup> in der HU-Stichprobe 2007 kann ein solcher (hypothetischer) Fahrleistungswert für die im Jahr 2007 nicht geprüften Pkw aus der Kohorte 2003 berechnet werden. Der Mittelwert 2007 der bisherigen Monatsfahrleistung *aller* Fahrzeuge aus der Pkw-Kohorte 2003 ergibt sich unter dieser Prämisse als gewogener Durchschnitt der (empirischen) mittleren bisherigen Monatsfahrleistung der 2007 geprüften und der (hypothetischen) mittleren bisherigen Monatsfahrleistung der 2007 nicht geprüften Fahrzeuge der Kohorte 2003.

Da in der HU-Stichprobe eines Jahres nur Fahrzeuge vertreten sind, die in diesem Jahr geprüft wurden, muss die oben skizzierte „übergreifende“ Fahrleistungsschätzung (für geprüfte und nicht geprüfte Fahrzeuge) datentechnisch so umgesetzt werden, dass die Stichprobe der geprüften Fahrzeuge die Gesamtheit der im betreffenden Jahr *geprüften und nicht geprüften* Fahrzeuge repräsentiert. Möglich wird dies durch eine geeignete Transformation der Variablen „mittlere Fahrleistung (km/Monat) seit Erstzulassung“.

#### 9.5.4 Verzerrungskorrektur nach dem Konzept der „neutralen“ Fahrleistung

Formal lässt sich die oben beschriebene Schätzung der mittleren Monatsfahrleistung seit Erstzulassung für das Bezugsjahr 2007 und die in diesem Bezugsjahr als „unregelmäßig“ einzustufende Kohorte  $e=2003$  als gewogener Durchschnitt aus einem empirischen und einem hypothetischen Fahrleistungswert<sup>41</sup> wie folgt darstellen:

$$(9-3) \quad s_e = s_e(g) \cdot P_e(g) + \frac{s_{e-1}(g) + s_{e+1}(g)}{2} \cdot P_e(u).$$

---

<sup>39</sup> Für die im Jahr 2007 *geprüften* Pkw der Kohorte 2003 sind keine Annahmen irgendwelcher Art notwendig. Da diese Fahrzeuge ja als „unregelmäßig“ geprüfte Fahrzeuge in der HU-Stichprobe 2007 vertreten sind, kann ihre mittlere bisherige Fahrleistung aus der HU-Stichprobe 2007 geschätzt werden.

<sup>40</sup> Die Kohorten 2002 und 2004 gehören im Bezugsjahr 2007 zu den „regelmäßig“ geprüften Pkw.

<sup>41</sup> Der hypothetischen Fahrleistungswert wird seinerseits natürlich aus zwei empirischen Werten berechnet

Hierbei bezeichnet  $s$  die bisherige Monatsfahrleistung (als Mittelwert über alle betreffenden Fahrzeuge in der HU-Stichprobe) und  $e$  das Zulassungsjahr. Das Symbol  $u$  steht für ungeprüfte Pkw,  $g$  für geprüfte Pkw und  $P$  ist der jeweilige Anteil in der Gesamtheit aller Pkw.

Der andere Fall - mittlere Monatsfahrleistung einer Kohorte, für die das Bezugsjahr ein „regelmäßiges“ Prüfwahl darstellt - ist analog zu behandeln. Im Beispiel des Bezugsjahres 2007 und der Fahrzeugart Pkw ist die angesprochene analoge Vorgehensweise aber erst ab der Kohorte 2004 möglich<sup>42</sup>, deren erstes „regelmäßiges“ Prüfwahl das Jahr 2007 ist.

Für die (drei) jüngeren Kohorten müssen deshalb weitere Annahmen getroffen werden. Eine mögliche Herangehensweise wäre im obigen Beispiel, die im Bezugsjahr 2007 ermittelte bisherige Monatsfahrleistung der Kohorte 2004 auch für die drei jüngeren Kohorten (2005, 2006, 2007) anzusetzen und damit zu unterstellen, dass die Fahrzeuge bis zu ihrer ersten „regelmäßigen“ Untersuchung eine konstante monatliche Fahrleistung erbringen.

Eine Analyse der Fahrleistungserhebung 2002 zeigt allerdings, dass die bisherigen monatlichen Fahrleistungen der jüngeren Kohorten im Vergleich zu den Pkw, die im Bezugsjahr 3 Jahre alt sind, höher liegen. Die Faktoren betragen: 1,22 für 0-jährige Fahrzeuge (hier Kohorte 2007); 1,19 für 1-jährige Fahrzeuge (hier Kohorte 2006) und 1,12 für 2-jährige Fahrzeuge (hier Kohorte 2005).

Für die älteren Kohorten sind vereinfachende Annahmen zu treffen. So werden bei den Pkw ab einem Alter von 20 Jahren die Fahrleistungen von geprüften und ungeprüften Fahrzeugen gleichgesetzt. Dies erfolgt einerseits, weil die Eigenschaft „regelmäßiges / unregelmäßiges Prüfwahl“ keinen Einfluss mehr auf die Fahrleistung nimmt, andererseits aus technischen Gründen, da nach diesem Fahrzeugalter (20 Jahre) die aggregierten Altersklassen folgen (vgl. Schichtungsplan in Abschnitt 9.4.4).

Die datentechnische Umsetzung des Korrekturverfahrens

„übergreifende Fahrleistungsschätzung für geprüfte und nicht geprüfte Fahrzeuge auf der Grundlage von Tachostandsdaten für geprüfte Fahrzeuge“

kann am Beispiel der Fahrzeugart Pkw für das Bezugsjahr  $t$  der Hochrechnung zusammenfassend wie folgt beschrieben werden:

---

<sup>42</sup> Nach der Kohorte 2004 folgen die Kohorten 2002, 2000, 1998 etc.



Die mittlere bisherige Monatsfahrleistung

$$s_a(u)$$

ungeprüfter Fahrzeuge der Altersgruppe  $a$  (entspricht der Kohorte  $e = t - a$ ) innerhalb einer Schicht (Hochrechnungsgruppe) wird in Abhängigkeit von der aus der HU-Stichprobe abgeleiteten mittleren bisherigen Monatsfahrleistung

$$s_e(g)$$

der geprüften Fahrzeuge berechnet:

$$s_0(u) = 1,22 \cdot s_3(g),$$

$$s_1(u) = 1,19 \cdot s_3(g)$$

$$s_2(u) = 1,12 \cdot s_3(g)$$

$$s_3(u) = \frac{s_2(g) + s_4(g)}{2}, \dots, s_{19}(u) = \frac{s_{18}(g) + s_{20}(g)}{2}$$

$$s_{20}(u) = s_{20}(g), \dots$$

Bei den übrigen Fahrzeugarten ist bei der Berechnung der hypothetischen Fahrleistung ungeprüfter Fahrzeuge analog zu verfahren, wobei aber die ebenfalls aus der Fahrleistungserhebung 2002 ermittelten Korrekturfaktoren für die 0- und 1-jährigen Fahrzeuge je nach Fahrzeugart unterschiedlich sind:

- Bei Lkw bis 3,5 t zulässigem Gesamtgewicht beträgt der Korrekturfaktor für die 0-jährigen Fahrzeuge 0,92 und für 1 Jahr alte Fahrzeuge 1,06 - jeweils bezogen auf die 2 Jahre alten Fahrzeuge.
- Für Lkw über 3,5 t zulässigem Gesamtgewicht gilt ein Korrekturfaktor von 0,87 für die 0 Jahre alten Fahrzeuge (hier bezogen auf die 1 Jahr alten Fahrzeuge).
- Bei Sattelzugmaschinen wird für 0 Jahre alte Fahrzeuge ein Korrekturfaktor von 0,92 angesetzt, auch hier bezogen auf die 1-jährigen Fahrzeuge.

Mit Hilfe der so berechneten hypothetischen mittleren Monatsfahrleistungen ungeprüfter Fahrzeuge wird in der Datenmatrix  $D_1$  - also auf Einzeldatenebene – für jedes Fahrzeug<sup>43</sup> eine neue Fahrleistungsvariable eingeführt, um Tabellierungen der hochgerechneten Jahresfahrleistung<sup>44</sup> nach beliebigen (in den HU-Daten enthaltenen) Gliederungsmerkmalen zu ermöglichen.

Die neue Fahrleistungsvariable  $\hat{s}_i$  berechnet sich für das i-te Fahrzeug aus einer beliebigen Schicht (Hochrechnungsgruppe) der HU-Stichprobe durch

$$(9-4) \quad \hat{s}_i = s_i \cdot P_{a(i)} + s_{a(i)}(u) \cdot [1 - P_{a(i)}] ,$$

wobei

$s_i$  die (empirische) bisherige Monatsfahrleistung des betrachteten Fahrzeugs  $i$ ,

$s_{a(i)}(u)$  die (hypothetische) mittlere bisherige Monatsfahrleistung ungeprüfter Fahrzeuge, die zur selben Schicht und Altersgruppe gehören wie das Fahrzeug  $i$  und

$P_{a(i)}$  den Grundgesamtheitsanteil<sup>45</sup> der geprüften Fahrzeuge innerhalb der betrachteten Fahrzeugschicht

bezeichnet.

Die neue Fahrleistungsvariable  $\hat{s}_i$  kann „neutrale“ oder auch „androgyn“ Monatsfahrleistung<sup>46</sup> genannt werden, da es sich hier um ein künstliches Merkmal der in der HU-Stichprobe enthaltenen geprüften Fahrzeuge handelt, welches es erlaubt, ausgehend von einer Stichprobe geprüfter Fahrzeuge für das betreffende Bezugsjahr die Jahresfahrleistung aller - im betreffenden Jahr geprüften und nicht geprüften - Fahrzeuge hochzurechnen.

---

<sup>43</sup> Die HU-Stichprobe enthält nur geprüfte Fahrzeuge.

<sup>44</sup> Die hochgerechnete Jahresfahrleistung bezieht sich auf die Gesamtheit *aller* Fahrzeuge, also der im betreffenden Jahr geprüften und nicht geprüften Fahrzeuge.

<sup>45</sup> Wie die betreffenden Anteilswerte berechnet werden, ist in Abschnitt 9.5.5 dargestellt.

<sup>46</sup> Die Bezeichnung „androgyn“ Fahrleistung wurde gewählt, da durch dieses künstliche Merkmal die geprüften Fahrzeuge aus der HU-Stichprobe zugleich auch die nicht geprüften Fahrzeuge aus dem gesamten ZFZR-Bestand repräsentieren und so also eine Art Zwitterstellung einnehmen.

### 9.5.5 Bestimmung der Anteilswerte geprüfter und ungeprüfter Fahrzeuge

In die Berechnung der „neutralen“ Fahrleistung eines Fahrzeugs aus der HU-Stichprobe fließen je Schicht die Anteilswerte der geprüften und ungeprüften Fahrzeuge bezogen auf die Grundgesamtheit aller Fahrzeuge ein. Näherungswerte für diese Anteile werden aus den mittleren ZFZR-Beständen unter Zuhilfenahme der FU-Statistik, die jährlich zur Verfügung steht, wie nachfolgend beschrieben berechnet – dabei wird davon ausgegangen, dass die Eckwerte der FU-Statistik (Gesamtzahl der durchgeführten Hauptuntersuchungen) differenziert nach Fahrzeugarten ausgewiesen werden.

Wenn man unterstellt, dass innerhalb einer Fahrzeugart die Aufteilung der Fahrzeuge in der HU-Stichprobe auf die einzelnen Schichten (Hochrechnungsgruppen) repräsentativ ist für die Gesamtheit aller Fahrzeuge der betreffenden Fahrzeugart, die im betrachteten Jahr einer HU unterzogen werden, lässt sich die *Gesamtzahl der durchgeführten Hauptuntersuchungen je Schicht* schätzen, indem man die Fallzahlen in der HU-Stichprobe auf den entsprechenden Eckwert gemäß FU-Statistik hochrechnet. Die im Rahmen der pilotartigen Anwendung des Hochrechnungsverfahrens verwendeten Gesamtzahlen von Hauptuntersuchungen sind am Beispiel des Untersuchungsjahres 2006 in der nachfolgenden Tabelle dargestellt:

**Tabelle 17: Gesamtzahl geprüfter Fahrzeuge im Untersuchungsjahr 2006**

Fahrzeugart	Gesamtzahl geprüfter Kfz
Krafträder	1.488.954
Pkw	17.856.821
Lkw	1.516.815
Sattelzugmaschinen	128.998
Sonstige Zugmaschinen	277.841
Zugmaschinen LoF	496.145
Omnibusse	24.676
Sonstige Kfz	87.486

Die benötigten Anteile  $P$  erhält man, indem für jede Schicht (Hochrechnungsgruppe) die nach der obigen Methode geschätzte Gesamtzahl geprüfter Fahrzeuge ins Verhältnis zum mittleren Fahrzeugbestand gesetzt wird.

Nach derzeitigem Stand sind die in einem Kalenderjahr tatsächlich untersuchten Kfz für die beiden Fahrzeugarten Pkw und Krafträder aus der FU-Statistik des KBA bekannt. Für die übrigen Fahrzeugarten liegt ein solcher Eckwert aber nur in aggregierter Form, d.h. als Summe über die betref-

fenden Fahrzeugarten, vor. Würde man aus der vorliegenden HU-Stichprobe die Verteilung der geprüften Fahrzeuge nach Fahrzeugart heranziehen und diese auf die für das genannte Aggregat (Gesamtzahl der Prüfungen aller „übrigen“ Fahrzeugarten) anwenden, so erhielte man unplausible Ergebnisse.

Beispielsweise wurden im Jahr 2007 rund 950 Tsd. Zugmaschinen einer HU unterzogen, in den vorliegenden HU-Daten machen die Sattelzugmaschinen rund 33 % aller Prüfungen von Zugmaschinen aus. Eine einfache Multiplikation dieses Anteils mit den Gesamtprüfzahlen würde also eine Anzahl von rund 313 Tsd. geprüften Sattelzugmaschinen ergeben. Im gesamten Kfz-Bestand befinden sich aber nur etwa 180.000 Fahrzeuge dieser Fahrzeugart.

Ein nicht zu unterschätzender Vorteil wäre es daher, wenn für die Hochrechnung einer kalenderspezifischen Gesamtfahrleistung die Prüfdaten aller Prüforganisationen vorlägen, da dann die benötigten Eckwerte (mittlerer Fahrzeugbestand des Untersuchungsjahres gegliedert nach geprüften und nicht geprüften Fahrzeugen) konkret berechnet werden könnten.

## **9.6 Hochrechnungsfaktoren: Gesamtzahl der Fahrzeugmonate und mittlerer Fahrzeugbestand im Bezugsjahr der Hochrechnung**

### **9.6.1 Bestimmung der Gesamtzahl der Fahrzeugmonate im Bezugsjahr**

Die Hochrechnung der Fahrleistung im Untersuchungsjahr erfolgt, indem die „neutrale“ mittlere bisherige Monatsfahrleistung  $\hat{s}_i$  der Fahrzeuge in der HU-Stichprobe mit dem zugehörigen Hochrechnungsfaktor  $g_i$  multiplikativ verknüpft wird. Der von der Schichtzugehörigkeit sowie dem Zulassungsjahr des Fahrzeugs abhängige Hochrechnungsfaktor ist dabei das Verhältnis zwischen der Gesamtzahl der Fahrzeugmonate des mittleren Fahrzeugbestands im Untersuchungsjahr und der entsprechenden Zahl der Fahrzeuge in der HU-Stichprobe. Die HU-Stichprobe wird so an die Struktur der Grundgesamtheit aller Fahrzeuge angepasst und auf die Gesamtzahl aller Fahrzeuge hochgerechnet.

Kfz-Bestandszahlen stehen jeweils für den 1.1. eines Jahres zur Verfügung. Im Rahmen des vorliegenden Hochrechnungsverfahrens stellt der Bestand zum 1.1. eines Jahres  $t$  zum einen den Anfangsbestand des Jahres  $t$  und zum anderen den Endbestand (31.12.) des Jahres  $t-1$  dar. Ausgehend von solchen Zahlen zum Anfangs- und Endbestand der Kfz ist für jedes Bezugsjahr der Hochrechnung die Gesamtzahl der Fahrzeugmonate zu berechnen.

Eine denkbare Herangehensweise bei der Bestimmung der Gesamtzahl der Fahrzeugmonate wäre, die genannten Bestandszahlen jeweils mit dem Faktor 12 zu multiplizieren. Ein solcher einfacher Ansatz kommt aus verschiedenen Gründen nicht in Betracht.

Der simple Ansatz „Fahrzeugmonate = Bestand x 12“ würde die Zahl der Fahrzeugmonate der im Bezugsjahr der Hochrechnung (das mit dem Prüfwahl der HU-Daten identisch ist) zugelassenen Fahrzeuge überschätzen, da hier unterstellt würde, dass alle diese Fahrzeuge im Januar des betreffenden Jahres erstangemeldet wurden. Da die Kfz-Bestandszahlen zum Ende eines Jahres auch gegliedert nach dem zu diesem Zeitpunkt erreichten Alter der Fahrzeuge (in Monaten) vorliegen, kann die Summe der Zulassungsmonate für die im Untersuchungsjahr erstmals zugelassenen Fahrzeuge eindeutig bestimmt werden. (Grundsätzlich wäre es sehr hilfreich, wenn beim KBA monatlich eine entsprechende Auszählung vorläge, da dann auch die Zulassungsmonate (seit Erstzulassung) von Fahrzeugen bestimmt werden könnten, die z.B. nur zwei Monate im aktiven Bestand waren.)

Für die älteren Fahrzeugkohorten gilt ebenfalls, dass sich deren Bestand im Laufe eines Kalenderjahres verändert. Zur näherungsweisen Kennzeichnung dieser Veränderung kann man getrennt nach Kohorten (Erstzulassungsjahr) den Endbestand des Untersuchungsjahres dem entsprechenden Anfangsbestand gegenüberstellen. Zur Illustration sind die Anfangs- und Endbestände 2007 für die Fahrzeugart „Sattelzugmaschinen“ in der folgenden Tabelle 18 dargestellt. Als durchschnittliche Bestände 2007 ließen sich nun einfache Mittelwerte berechnen; für die Kohorte 2006 (Erstzulassungsjahr) ergäbe sich als mittlerer Bestand des Jahres 2007 der folgende Wert:  $(33.128+32.383)/2 = 32.756$  Kfz. Diese Bestandszahl könnte man mit 12 multiplizieren, um die Zahl der Fahrzeugmonate zu erhalten.

Bei diesem Ansatz würde allerdings unterstellt, dass Fahrzeuge einer Kohorte, die im Untersuchungsjahr (hier 2007) aus dem Bestand ausscheiden, durchschnittlich noch über 6 Monate hinweg Fahrleistungen erbringen, was etwas zu hoch erscheint. Aus Plausibilitätsüberlegungen wurden diesen Fahrzeugen nur 3 Fahrzeugmonate für Untersuchungsjahr zugeordnet.<sup>47</sup> Aus pragmatischen Gründen wurden die jeweils letzten beiden aggregierten Kohorten gemittelt. Hinzuweisen ist noch darauf, dass bei nicht bekanntem Alter diese Fahrzeuge jeweils der höchsten Alterskategorie zugeordnet wurden.

---

<sup>47</sup> Dies beruht auf der Überlegung, dass der Ersatz von Fahrzeugen primär im Frühjahr erfolgt. Es sind hier aber natürlich auch andere Annahmen denkbar, die problemlos in das Hochrechnungsverfahren einfließen könnten.

**Tabelle 18: Bestandszahlen und Fahrzeugmonate von Sattelzugmaschinen 2007 nach Zulassungsjahren**

Bestandsauszählung 31.12.2006 (Anfangsbestand 2007)		Bestandsauszählung 31.12.2007 (Endbestand 2007)		Kfz-Monate im Jahr 2007 für die Hoch- rechnung	Mittlerer Kfz- Bestand 2007 für die Hoch- rechnung
Zulassungsjahr	Bestand	Zulassungsjahr	Bestand		
-	-	2007	36188	240273	36188
2006	33128	2006	32383	390831	32569
2005	29372	2005	27554	336102	28009
2004	25818	2004	22172	277002	23084
2003	18364	2003	14323	183999	15333
2002	12798	2002	10080	129114	10760
2001	10959	2001	8655	110772	9231
2000	9682	2000	7487	96429	8036
1999	7945	1999	6184	79491	6624
1998	5228	1998	4043	52071	4339
1997	3058	1997	2367	30477	2540
1996	1932	1996-1987	7543	90078	7507
1995 - 1986	7470	1986 und früher	956	11148	929
1985 und früher	902	-	-	-	-

Konkret wird die Gesamtzahl der Fahrzeugmonate aller zur Grundgesamtheit gehörenden Fahrzeuge in einem bestimmten Untersuchungsjahr wie folgt berechnet:

$$\text{Anzahl Fahrzeugmonate} = (\text{Endbestand}) \times 12 + (\text{Anfangsbestand} - \text{Endbestand}) \times 3$$

In diesem Zusammenhang ist zu beobachten, dass es ganz vereinzelt auch Schichten gibt, bei denen der Jahresendbestand einer Kohorte größer als der Jahresanfangsbestand der betreffenden Kohorte ist. In diesem Fall (Endbestand der Kohorte größer als Anfangsbestand) wird die Gesamtzahl der Fahrzeugmonate wie folgt berechnet:

$$\text{Anzahl Fahrzeugmonate} = (\text{Anfangsbestand}) \times 12 + (\text{Endbestand} - \text{Anfangsbestand}) \times 3$$

### 9.6.2 Schätzung des mittleren Fahrzeugbestands im Bezugsjahr

Für die tabellarische und grafische Darstellung von Fahrleistungskennzahlen wird in vielen Fällen auch der mittlere Fahrzeugbestand des betreffenden Untersuchungsjahres benötigt:

- Bei Tabellierungen des Jahrestotalwertes der Fahrleistung ist es zweckmäßig, zusätzlich den zugehörigen Fahrzeugbestand anzugeben; dadurch erhält man eine zahlenmäßige Vorstellung von der Größe des Fahrzeugkollektivs, welches den betreffenden Fahrleistungstotalwert „hervorgebracht“ hat.
- In manchen Zusammenhängen interessiert die „mittlere Jahresfahrleistung“ (km pro Fahrzeug und Jahr) mehr als der Totalwert der Jahresfahrleistung; die Mittelwertberechnung setzt naturgemäß die Kenntnis des Fahrzeugbestands voraus.

Ein Schätzwert für den mittleren Fahrzeugbestand im Untersuchungsjahr wird schließlich je Schicht (Hochrechnungsgruppe) gemäß der Formel

$$\text{Mittlerer Fahrzeugbestand} = (\text{Gesamtzahl der Fahrzeugmonate})/12$$

berechnet.

Die aus dieser Vorgehensweise für das Untersuchungsjahr 2007 resultierenden jahresdurchschnittlichen Bestände von Sattelzugmaschinen sind untergliedert nach Kohorten in der vorletzten Spalte der Tabelle 18 dargestellt, während in der letzten Spalte die Gesamtzahl der Fahrzeugmonate im Untersuchungsjahr 2007, dem Bezugsjahr der Hochrechnung, enthält.

## 9.7 Grenzen des Hochrechnungsverfahrens

Die Möglichkeiten und Grenzen des Hochrechnungsverfahrens werden in Kapitel 11 zusammenfassend dargestellt. Vorab sei aber bereits an dieser Stelle darauf hingewiesen, dass das hier entwickelte Durchschnittsverfahren Strukturbrüche der Fahrleistungsentwicklung am „aktuellen Rand“ - etwa als Folge drastisch gestiegener Kraftstoffpreise - nur teilweise diagnostizieren kann.

Verändern sich die Rahmenbedingungen des Kraftfahrzeugverkehrs in der aktuellen Periode gravierend, so hat dies Auswirkungen auf

- die Neuzulassung sowie die Abmeldung von Kraftfahrzeugen und damit auf den mittleren Fahrzeugbestand und dessen Struktur,
- die vorübergehende Stilllegung von Fahrzeugen und damit auf die mittlere Anmeldequote (speziell im Bereich der Nutzfahrzeuge) sowie schließlich auf

- die Nutzungsintensität der angemeldeten Fahrzeuge gemessen durch die mittlere Tages- bzw. Monatsfahrleistung.

Veränderungen im Totalwert der Fahrleistungen, die auf Veränderungen von Zahl, Art und Anmeldestatus der Fahrzeuge zurückgehen, werden durch das Hochrechnungsverfahren direkt erfasst, da ja die im ZFZR als „angemeldet“ registrierten Fahrzeugen den Hochrechnungsrahmen bilden. Soweit also z.B. gewerbliche Kfz-Halter auf gestiegene Kraftstoffpreise (oder Konjunkturschwächen) mit häufigerer vorübergehender Stilllegung von Fahrzeugen reagieren, werden die damit verbundenen Fahrleistungseffekte im Hochrechnungsergebnis widergespiegelt.

Anders stellt sich dies im Hinblick auf sprunghafte Veränderungen in der Nutzungsintensität der angemeldeten Fahrzeuge dar. Geht man in einem hypothetischen Beispiel davon aus, dass es zwischen 2006 und 2007 eine Erhöhung der Kraftstoffpreise um 50 % gegeben hätte, so hätte dies bei der Fahrzeuggruppe Pkw zu einem Rückgang der Jahresfahrleistung um etwa 15 % geführt (Elastizität der Nachfrage -0,3; vgl. Hautzinger u.a. 2004). Die Fahrleistung wäre demnach vom Normalniveau 100 % auf das Niveau 85 % gesunken. Anstelle des Totalwerts 586,2 Mrd. km (vgl. Tabelle 19) hätte sich dann für 2007 der Wert 498,3 Mrd. km ergeben.

Mit dem Durchschnittswertverfahren wäre der Rückgang (hier wird als Extremfall angenommen, dass der durch die 50-prozentige Kraftstoffpreissteigerung verursachte Fahrleistungsrückgang um 15 % *allein* aus der gesunkenen Nutzungsintensität der angemeldeten Fahrzeuge resultiert!) unterschätzt worden. Bei einem Durchschnittsalter der Fahrzeuge von 8 Jahren hätte sich Folgendes ergeben:

Wegen  $(100 \times 7 + 85 \cdot 1) / 8 = 98,125$  wäre mit dem Durchschnittswertverfahren das Fahrleistungsniveau für die aktuelle Bezugsperiode 2007 in erster Näherung nicht auf 85 % sondern auf 98 % geschätzt worden. Das Hochrechnungsergebnis für 2007 hätte demnach 574,5 Mrd. km betragen ( $586,2 \times 0,98 = 574,5$ ) und wäre damit deutlich zu hoch ausgefallen.

Welche Konsequenzen die begrenzte Empfindlichkeit des Hochrechnungsverfahrens gegenüber Strukturbrüchen am aktuellen Rand des Untersuchungszeitraums für den Einsatz des Verfahrens hat, wird wie erwähnt in Kapitel 11 behandelt.



## 10 Hochrechnungsergebnisse

### 10.1 Hochgerechnete Jahresfahrleistungen 2006 und 2007

Mit Hilfe des in Kapitel 9 beschriebenen statistischen Verfahren wurde aus den HU-Stichproben der Jahre 2006 und 2007 die Jahrleistung 2006 und 2007 der in Deutschland zugelassenen Kraftfahrzeuge (Inländerfahrleistung hochgerechnet. In Tabelle 19 sind die hochgerechneten Jahrestotalwerte der Fahrleistung untergliedert nach den hier betrachteten Fahrzeugarten aufgelistet und den veröffentlichten Ergebnissen der DIW-Modellrechnungen gegenübergestellt. Wie man sieht, liegt der auf der Hochrechnung von HU-Daten basierende *Schätzwert der Jahresfahrleistung* aller in der Bundesrepublik Deutschland zugelassenen Kfz bei

671,5 Mrd. km für das Untersuchungsjahr 2006

und bei

679,3 Mrd. km für das Untersuchungsjahr 2007.

Mit diesen Zahlen stehen neue Eckdaten zur Inländerfahrleistung zur Verfügung. Es liegt nahe, diese neuen verkehrstatistischen Informationen mit bereits bestehenden Datenquellen zu vergleichen. Da die letzte spezielle Fahrleistungserhebung 2002 inzwischen schon einige Jahre zurückliegt, kommt im Wesentlichen nur ein Vergleich mit der DIW-Modellrechnung in Betracht.

### 10.2 Vergleich mit Ergebnissen der DIW-Modellrechnungen

#### 10.2.1 Grad der Übereinstimmung

Betrachtet man die Ergebnisse für die Untersuchungsjahre 2006 und 2007, so zeigt sich in der Summe über alle Fahrzeugarten eine hohe Übereinstimmung zwischen den aus den HU-Daten hochgerechneten Fahrleistungen (671,5 bzw. 679,3 Mrd. km) und den nach einem vollkommen unterschiedlichen Ansatz<sup>48</sup> ermittelten DIW-Ergebnissen (682,7 bzw. 687,4 Mrd. km). Die relative Abweichung der auf HU-Daten basierenden Fahrleistungsschätzung vom DIW-Wert liegt für 2006 bei -1,6 % und für 2007 bei -1,2 %.

---

<sup>48</sup> Ergänzend sei hier auch nochmals darauf hingewiesen, dass es bei der Fahrleistungsschätzung auf Basis von HU-Daten keine wie auch immer geartete Anpassung an DIW-Kennzahlen gibt. Einzig die verwendeten Eckwerte des Kfz-Bestands sind bei beiden Schätzverfahren dieselben.

**Tabelle 19: Hochrechnungsergebnisse zur Jahresfahrleistung 2006 und 2007**

Fahrzeugart	DIW-Ergebnis* zu Fahrleistungen (Mio. km)		Fahrleistungsschätzung aus HU-Daten (Mio. km)		Schätzung aus HU-Daten bezogen auf DIW-Ergebnis		
	Jahr	2006	2007	2006	2007	2006	2007
Krafträder		13.213	10.841	9.713	9.604	73,5 %	88,6 %
Pkw gesamt		583.905	587.543	581.319	586.195	99,6 %	99,8 %
- darunter Benzin		378.705	370.696	372.580	363.252	98,9 %	98,0 %
- darunter Diesel		205.200	216.846	208.245	222.024	100,1 %	102,4 %
Omnibusse		3.502	3.402	4.047	4.083	115,6 %	120,0 %
Lastkraftwagen		57.649	59.845	50.557	52.617	87,7 %	87,9 %
Sattelzugmaschinen		16.604	17.801	18.036	19.339	108,6 %	108,6 %
Restliche Zugmaschinen		4.281	4.414	3.982	3.711	93,0 %	84,1 %
Sonstige Zugmaschinen		-	-	3.143	2.890	-	-
Zugmaschinen LoF		-	-	839	820	-	-
Sonstige Kfz		3.568	3.603	3.817	3.730	107,0 %	103,5 %
Gesamt		682.722	687.449	671.472	679.279	98,4 %	98,8 %

\* Quelle: DIW Wochenbericht Kraftfahrzeugverkehr 2007, S. 796 ff.

Die weitgehende Übereinstimmung der Gesamtwerte der Jahresfahrleistung (Hochrechnung HU-Daten versus DIW-Kraftstoffverbrauchsrückrechnung) resultiert vor allem aus der extrem guten Übereinstimmung bei der zahlenmäßig bedeutendsten Fahrzeugart Pkw, wo die Abweichung nur bei -0,4 % bzw. -0,2 % liegt und auch bei einer zusätzlichen Aufgliederung nach der Antriebsart des Fahrzeugs (Benzin, Diesel) nur geringe Ergebnisunterschiede zu beobachten sind.

Bei den Sattelzugmaschinen und den Sonstigen Kfz betragen die Ergebnisunterschiede zwischen Hochrechnung und DIW-Modellrechnung weniger als 10 %, was auch noch als relativ gute Übereinstimmung gesehen werden kann. Größere Abweichungen (über 10 %) findet man für die Fahrzeugarten Lkw, Omnibusse und Krafträder. Bei den Restlichen Zugmaschinen liegt die Abweichung 2006 unter und die Abweichung 2007 über 10 %.

Anmerkungen zu Tabelle 19: Bei der Fahrzeugart „Krafträder“ konnte nicht der gesamte gezählte Bestand 2007 von 3.566.122 Fahrzeugen für die Hochrechnung herangezogen werden, da bei einigen Krafträdern sowohl die Altersangabe als auch das Fahrzeugmerkmal „KW“ fehlt. Die Anzahl solcher Krafträder im Bestand ist mit 14.375 Fahrzeugen im Vergleich zum Gesamtbestand allerdings zu vernachlässigen.

Bei der Fahrzeugart „Pkw“ sind im Hochrechnungsrahmen für „Pkw mit Hybrid/sonstigem Antrieb“ auch Fahrzeuge ohne Angabe zur Antriebsart enthalten. Deren Zahl ist allerdings ebenfalls sehr gering und liegt bei 610 Pkw.

### **10.2.2 Mögliche Ursachen für Ergebnisunterschiede**

Die Abweichung der Fahrleistungskennzahlen für die Fahrzeugart Omnibusse (+15,6 % bzw. +20,0 %) kann auf der Tatsache basieren, dass in den HU-Daten keine Linienbusse enthalten sind (vgl. „Hochrechnung der Kfz-Fahrleistung auf der Basis der Ablesungen der Wegstreckenzähler bei Kfz-Hauptuntersuchungen“, Projekt-Nummer 70.0800/2006/, S 32.), so dass, wenn Linienbusse gegenüber den anderen Bussen eine geringere Fahrleistung aufweisen, die Gesamtfahrleistung von Omnibussen mit den hier verwendeten Daten überschätzt wird (im Hochrechnungsrahmen sind alle Omnibusse enthalten).

Relativ groß sind auch die Ergebnisunterschiede bei der Fahrzeugart Lkw (-12,3 % bzw. -12,1 %). Hierfür können aus Sicht der vorliegenden Studie keine plausiblen Erklärungen angeboten werden<sup>49</sup>. Die aus den HU-Daten ermittelten durchschnittlichen Lkw-Jahresfahrleistungen 2006 und 2007 stimmen im Hinblick auf die Fahrleistung nach Altersklassen (vgl. Abbildung 14) strukturell gut mit den Resultaten der Fahrleistungserhebung 2002 überein.

Bei Sattelzugmaschinen erhält man durch Hochrechnung der HU-Daten eine gegenüber den Modellrechnungen des DIW höhere Fahrleistung (jeweils +8,6 %). Entsprechende Analysen der HU-Daten zeigen aber keine Auffälligkeiten; verdeutlicht wird dies durch Abbildung 15, wo die Ergebnisse der Hochrechnung der HU-Daten für beide Bezugsjahre nach Fahrzeugalter dargestellt sind.

Bei der Fahrzeugart „Sonstige Kfz“ (Abweichung +7,0 % bzw. + 3,5 %) ist zu beachten, dass die in diesem Projekt vorgenommene Zuordnung der Fahrzeuge zu Fahrzeugarten für die Fahrzeugart „Sonstige Kfz“ einen höheren Bestand ergibt, als dies in den DIW-Modellrechnungen der Fall ist. Die DIW-Rechnungen weisen einen Durchschnittsbestand für das Jahr 2007 von 2,59 Mio. Kfz aus, der für die Hochrechnung der HU-Daten verwendete mittlere Bestand beträgt 2,71 Mio. Kfz, liegt also rund 5 % höher.

---

<sup>49</sup> Vergleichende Genauigkeitsbewertungen (HU-Daten versus DIW-Ergebnisse) wären letztlich nur mit einer neuen Fahrleistungserhebung nach dem Muster der FLE 2002 möglich.

## 10.3 Beispiele zur Ergebnisdarstellung

### 10.3.1 Jahresfahrleistung von Teilgruppen unterschiedlicher Fahrzeugarten

Die geschätzten Totalwerte der Jahresfahrleistung können innerhalb der Fahrzeugarten nach weiteren Fahrzeugmerkmalen aufgegliedert werden. Die nachfolgenden drei Tabellen zeigen beispielartig Darstellungsmöglichkeiten für Krafträder, Pkw und Lkw.

**Tabelle 20: Jahresfahrleistung 2006 von Krafträdern gegliedert nach Motorleistung**

Motorleistung in KW	Jahresfahrleistung 2006 in Mio. km	Jahresfahrleistung 2006 in %
bis 15 KW	2.141,1	22,0
16 bis 64 KW	4.551,2	46,9
65 KW und mehr	3.020,3	31,1
Insgesamt	9.712,6	100,0

**Tabelle 21: Jahresfahrleistung 2006 von Pkw gegliedert nach Antriebsart und Motorleistung**

Antriebsart und Motorleistung in KW	Jahresfahrleistung 2006 in Mio. km	Jahresfahrleistung 2006 in %
Pkw Benzin		
- bis 44 KW	63.912,3	11,0
- 45 bis 59 KW	71.623,0	12,3
- 60 bis 89 KW	135.913,6	23,4
- 90 KW und mehr	99.856,2	17,2
Pkw Diesel		
- bis 44 KW	3.105,2	0,5
- 45 bis 59 KW	18.913,9	3,3
- 60 bis 89 KW	94.464,2	16,2
- 90 KW und mehr	91.762,0	15,8
Pkw Hybrid/sonstiger Antrieb	1.768,7	0,3
Insgesamt	581.319,1	100,0

**Tabelle 22: Jahresfahrleistung 2006 von Lkw gegliedert nach Antriebsart bzw. zulässigem Gesamtgewicht**

Gliederungsmerkmal	Jahresfahrleistung 2006 in Mio. km	Jahresfahrleistung 2006 in %
<b>Antriebsart</b>		
- Benzin	1.960,5	3,9
- Diesel	48.568,6	96,1
- Sonstige/Unbekannt	27,8	0,1
<b>Insgesamt</b>	<b>55.556,9</b>	<b>100,0</b>
<b>Zulässiges Gesamtgewicht</b>		
- bis 3,5 t	31.169,4	61,7
- über 3,5 t bis 7,5 t	6.313,0	12,5
- über 7,5 t	13.074,5	25,9
<b>Insgesamt</b>	<b>55.556,9</b>	<b>100,0</b>

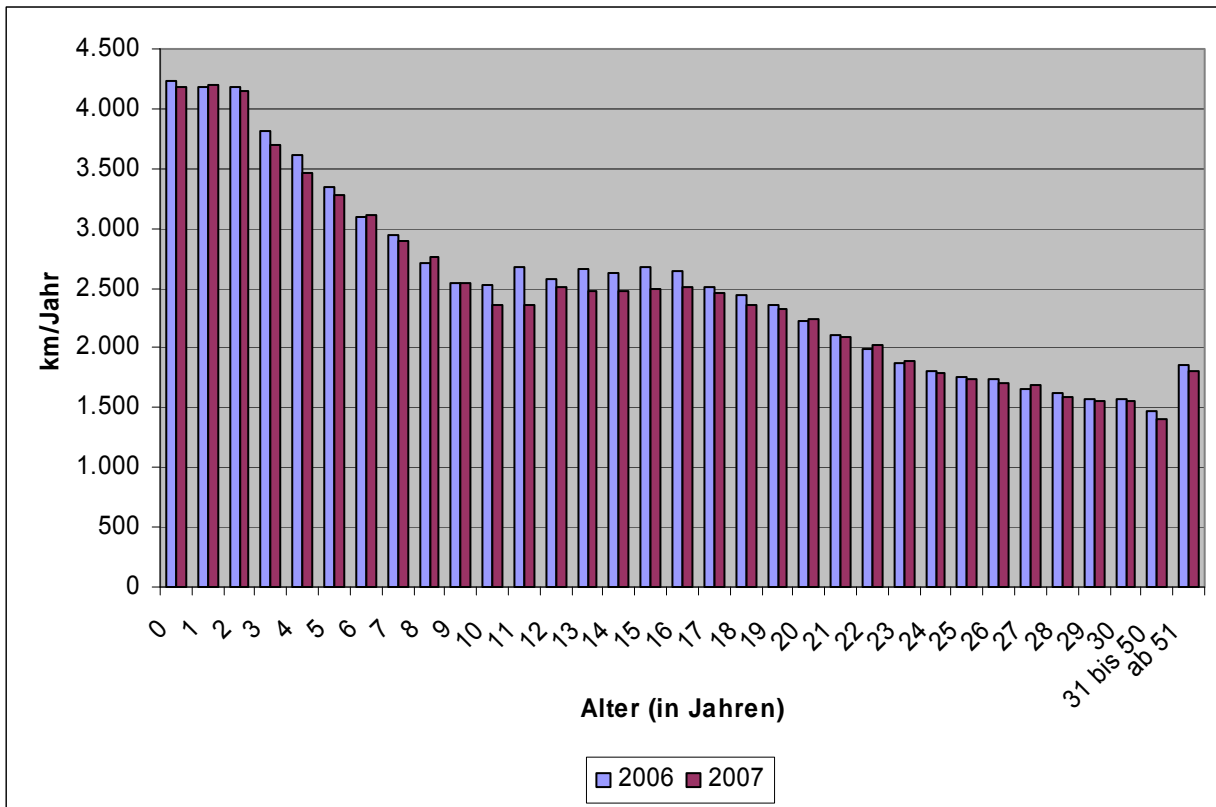
Zur Aufgliederung der Jahresfahrleistung 2006 von Lkw nach dem Merkmal „zulässiges Gesamtgewicht“ sei hier ergänzend angemerkt, dass die Anteilswerte der drei Größenklassen recht gut mit den entsprechenden Anteilswerten gemäß FLE 2002 übereinstimmen, die 57,3 bzw. 15,9 bzw. 26,8 % betragen. Das Gesamttotal 2006 für Lkw liegt mit 55,6 Mrd. km um rund 8 % unter dem Totalwert 2002 gemäß FLE (60,2 Mrd. km).

### 10.3.2 Mittlere Jahresfahrleistungen 2006 und 2007 nach Fahrzeugart und -alter

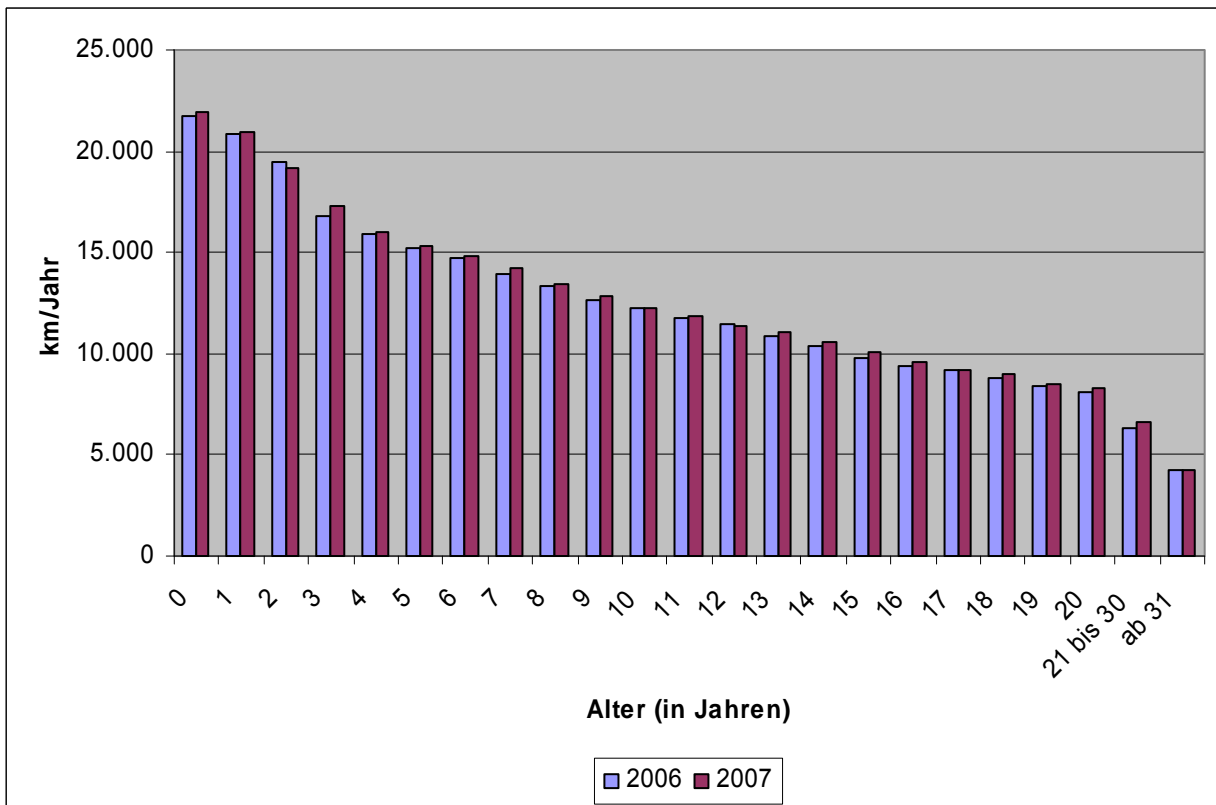
Wie bereits ausgeführt, können die aus HU-Daten hochgerechneten Jahresfahrleistungen prinzipiell nach (fast) allen im HU-Datensatz enthaltenen Merkmalen aufgegliedert werden. Bei der Tabellierung und grafischen Darstellung von Hochrechnungsergebnissen bieten sich vor allem die Gesamtfahrleistung in Mio. km pro Jahr und die mittlere Jahresfahrleistung in km pro Fahrzeug und Jahr als Fahrleistungskennzahlen an. Bei der Ergebnisdarstellung wird man nahezu immer nach Fahrzeugarten unterscheiden.

Als Beispiel können die folgenden Abbildungen dienen, welche für die verschiedenen Fahrzeugarten die mittleren Jahresfahrleistungen 2006 und 2007 (Gesamtfahrleistung bezogen auf den mittleren Fahrzeugbestand) in Abhängigkeit vom Gliederungsmerkmal Alter des Fahrzeugs zeigen.

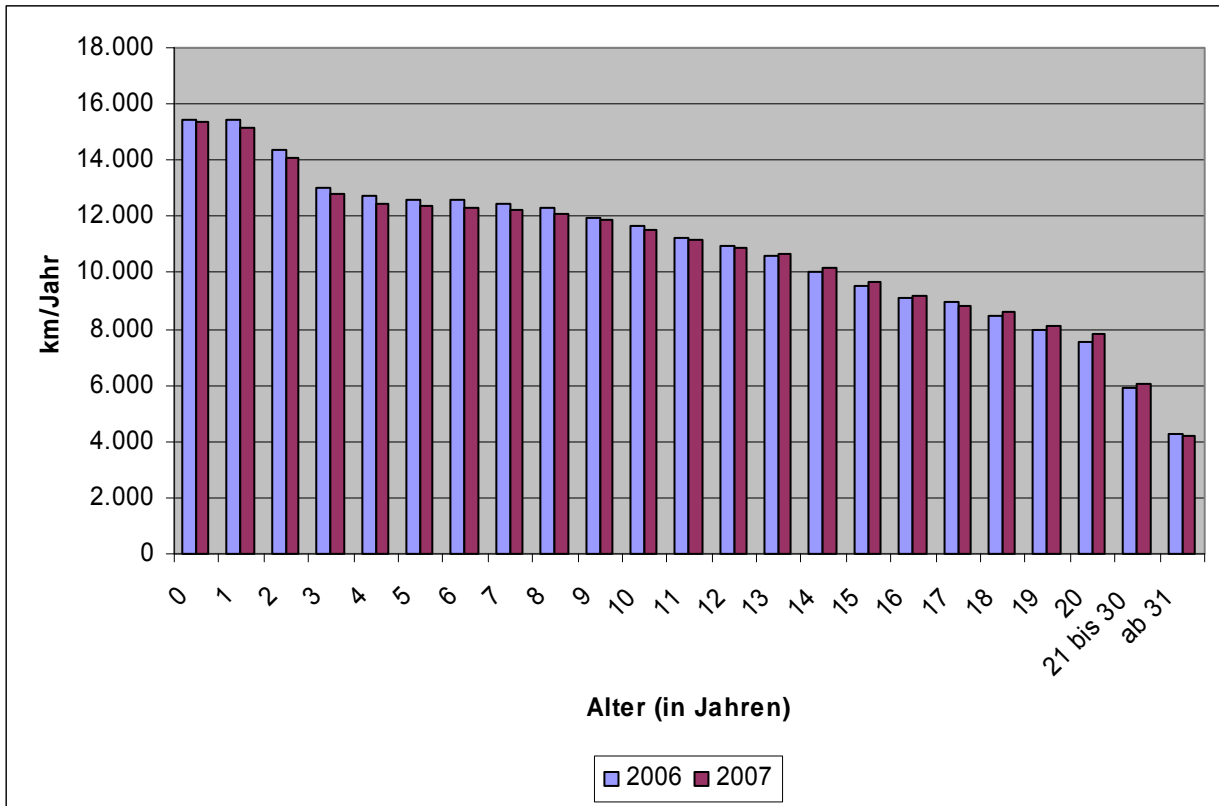
**Abbildung 10** Mittlere Jahresfahrleistung Krafträder



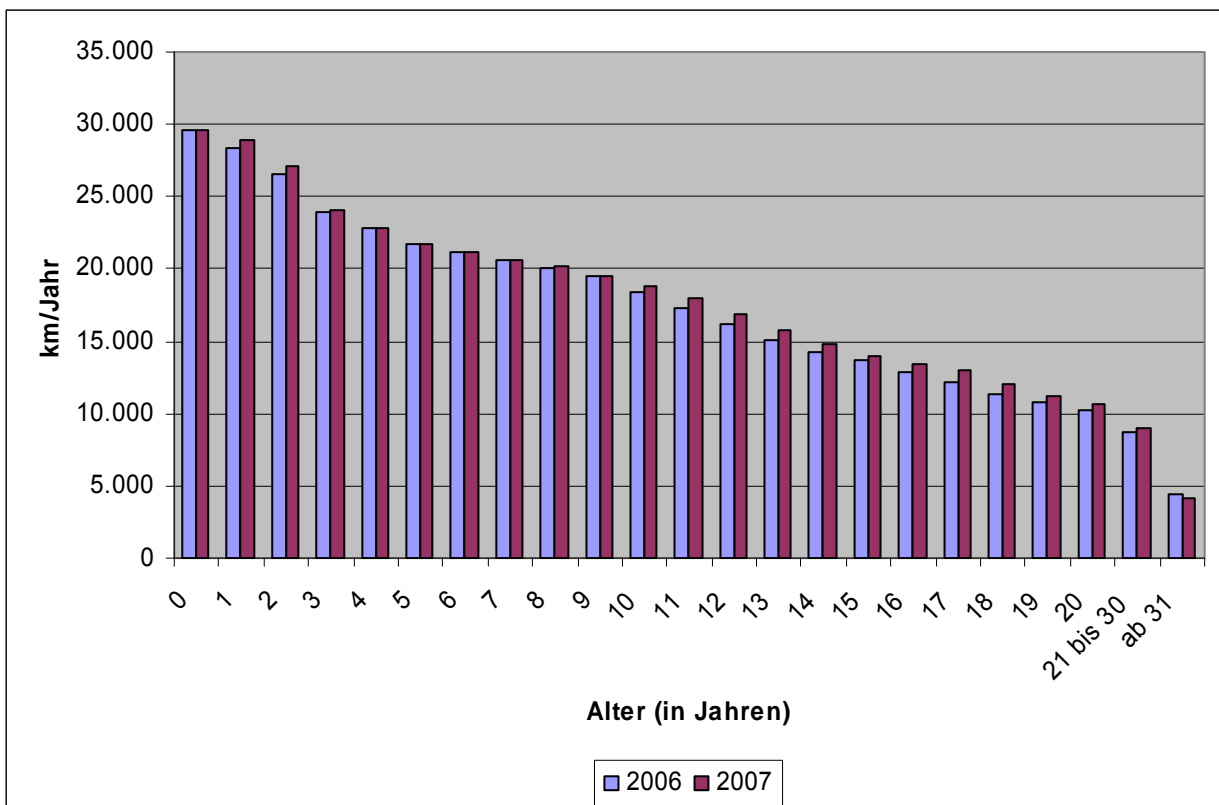
**Abbildung 11** Mittlere Jahresfahrleistung Pkw



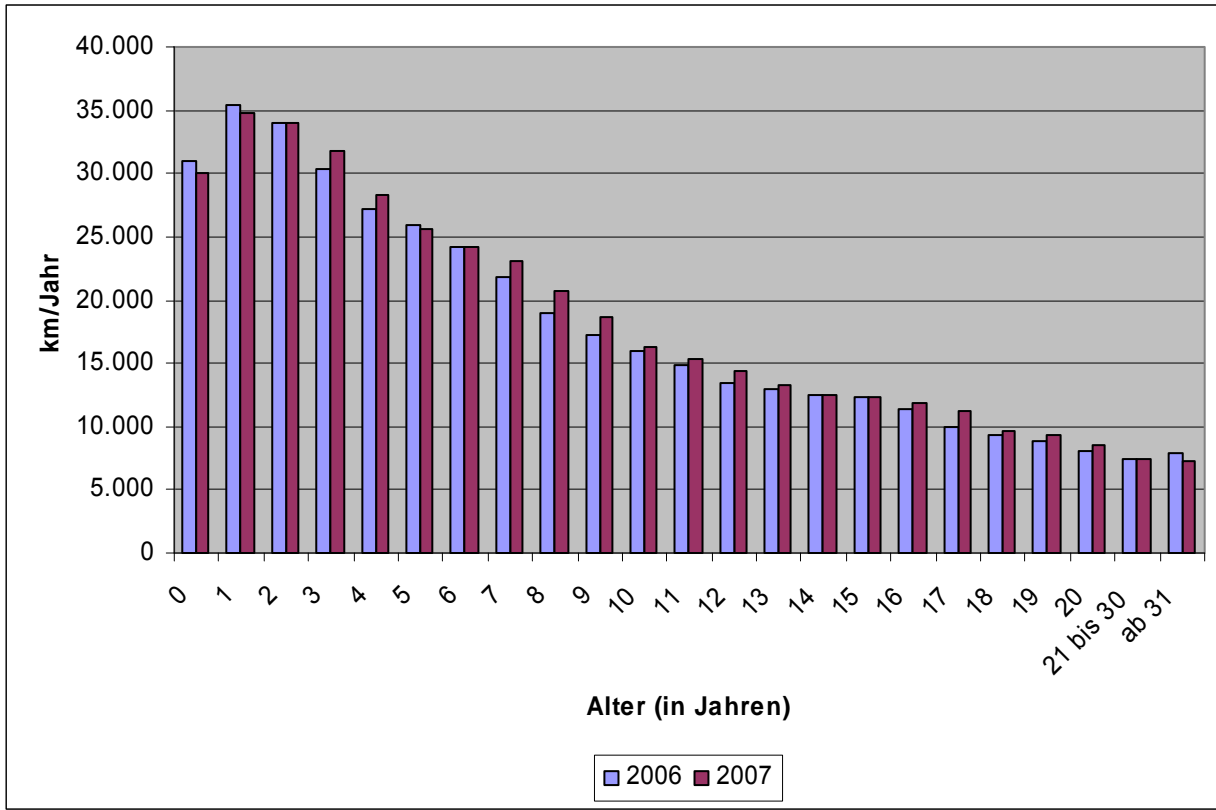
**Abbildung 12** Mittlere Jahresfahrleistung Pkw (Benzin)



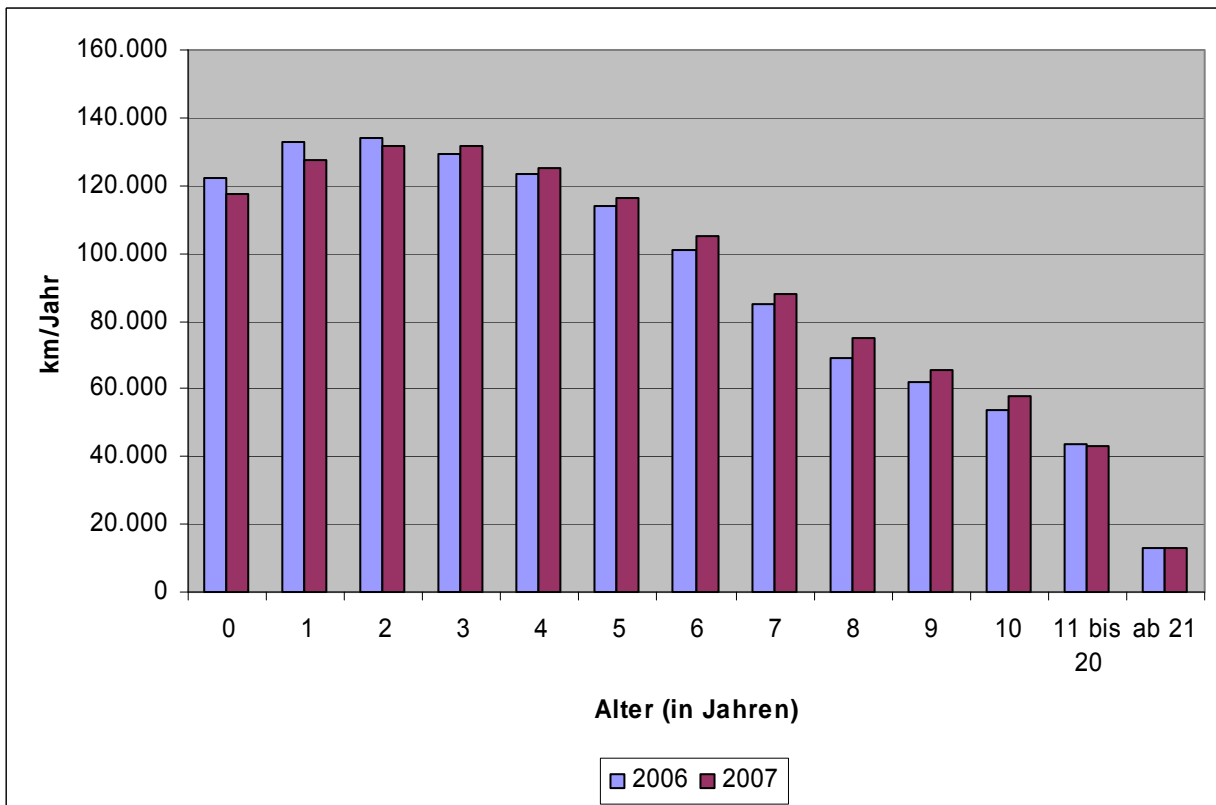
**Abbildung 13** Mittlere Jahresfahrleistung Pkw (Diesel)



**Abbildung 14** Mittlere Jahresfahrleistung Lkw

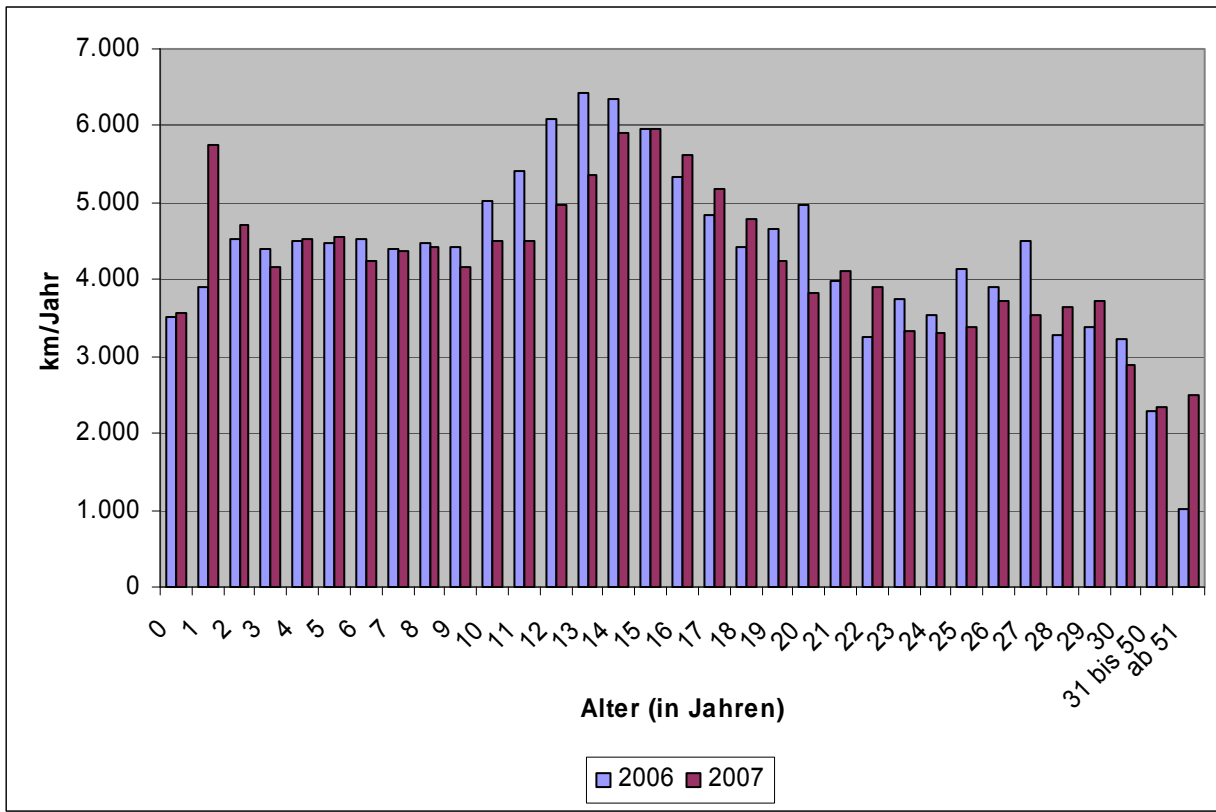


**Abbildung 15** Mittlere Jahresfahrleistung Sattelzugmaschinen

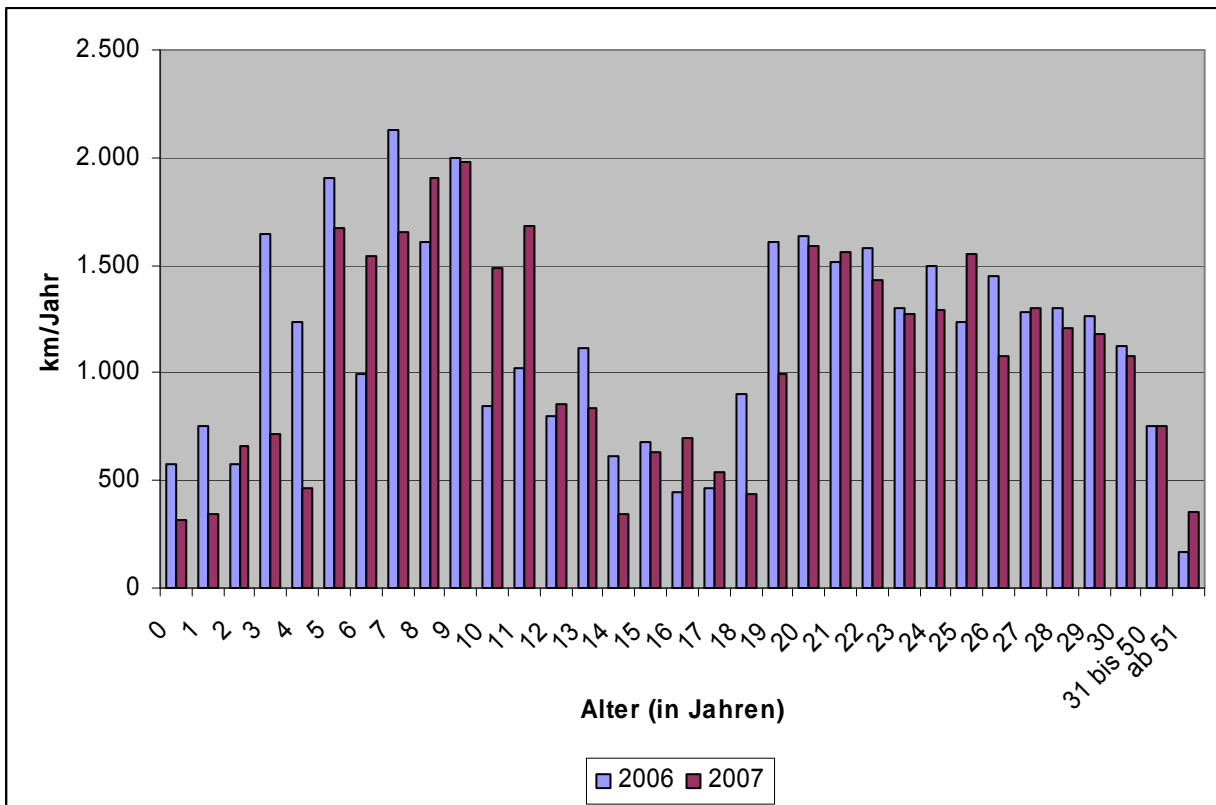




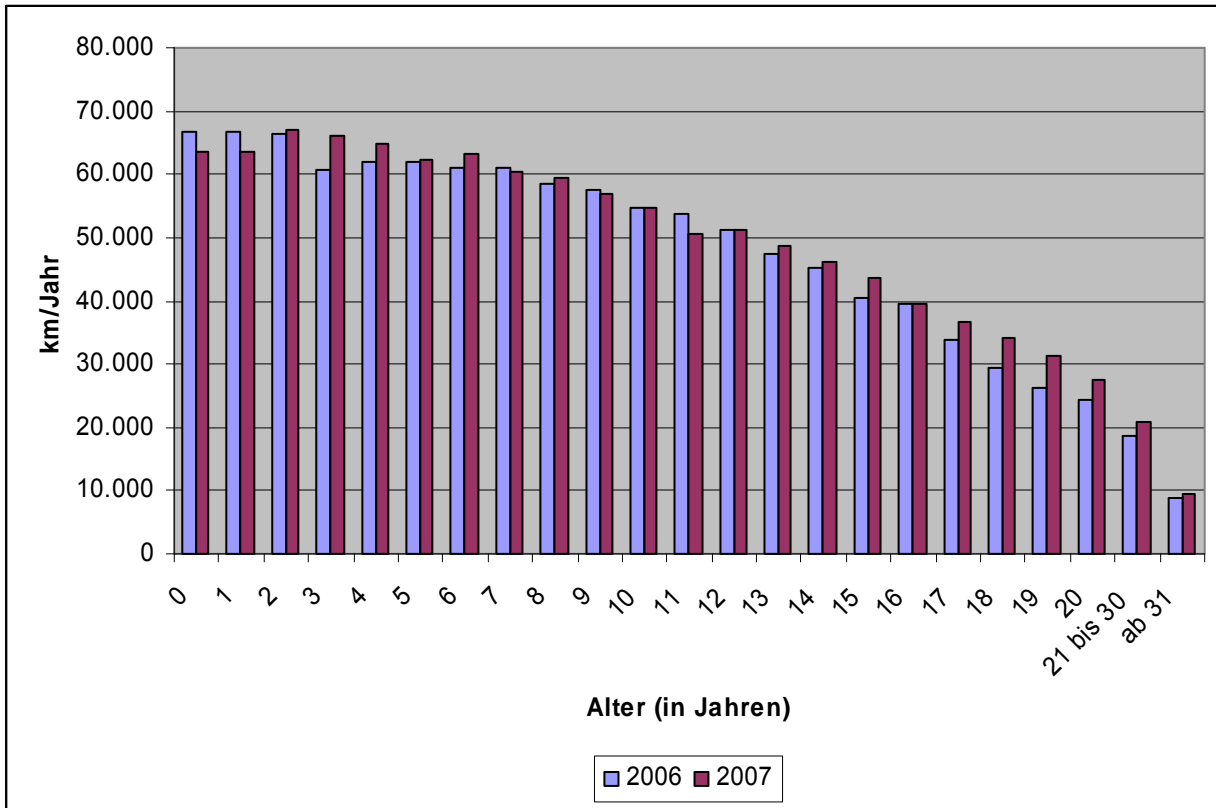
**Abbildung 16** Mittlere Jahresfahrleistung Sonstige Zugmaschinen



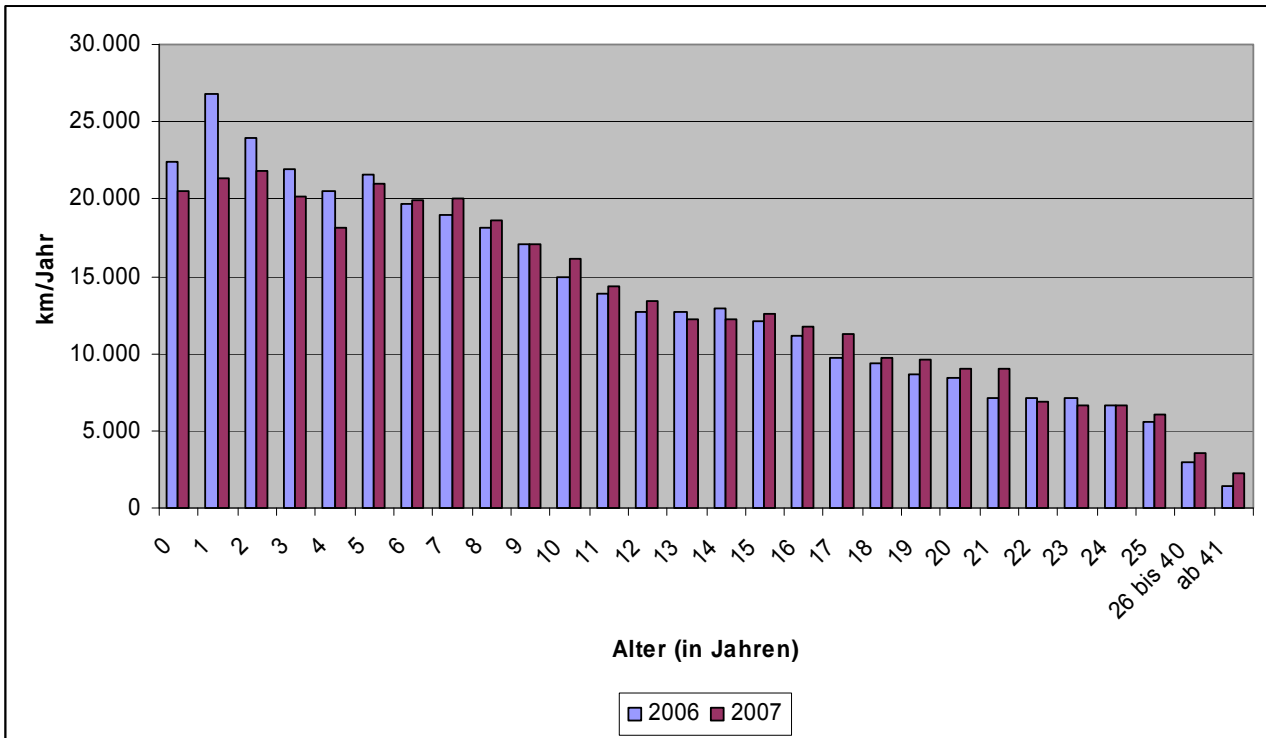
**Abbildung 17** Mittlere Jahresfahrleistung Zugmaschinen LoF



**Abbildung 18** Mittlere Jahresfahrleistung Omnibusse



**Abbildung 19** Mittlere Jahresfahrleistung Sonstige Kfz



## 11 Kritische Diskussion des entwickelten Hochrechnungsverfahrens

Im Rahmen des vorliegenden Forschungsvorhabens sollte ein Verfahren entwickelt und getestet werden, welches als methodischer Kern einer neu zu etablierenden laufenden Fahrleistungsstatistik dazu geeignet ist, auf der Basis von Daten aus Kfz-Hauptuntersuchungen kalenderjahrbezogene Schätzwerte für die Fahrleistung der in der Bundesrepublik Deutschland zugelassenen Kraftfahrzeuge zu liefern. Ein solches Verfahren, bestehend aus den beiden Komponenten „Aufbereitung der HU-Daten“ und „Hochrechnung der HU-Daten“ liegt als Resultat intensiver Forschungs- und Entwicklungsarbeiten nunmehr vor.

Wie ist das Ergebnis der Methodenentwicklung aus Sicht der Entwickler nun zu bewerten?

Einer solchen Bewertung sind einige Feststellungen vor auszuschicken:

- Für die Hochrechnung periodenbezogener Fahrleistungen sollten naturgemäß Daten über Fahrleistungen in der fraglichen Periode die empirische Grundlage bilden, nicht Daten zum Tachostand von Fahrzeugen an einem Stichtag innerhalb der Periode.
- Die naheliegendste Verwendung von Tachostandsdaten ist ihre Nutzung zur Schätzung des mittleren Tachostands und damit der mittleren bisherigen Gesamtfahrleistung von Fahrzeugen. Der mittlere Tachostand als Kennzahl für den „Abnutzungs- bzw. Alterungsgrad“ einer Fahrzeugflotte ist in der amtlichen und halbamtlichen Verkehrsstatistik aber nur von geringem Interesse.
- Dass man Tachostandsdaten trotzdem zur Schätzung periodenbezogener Fahrleistungen heranziehen möchte, liegt daran, dass sie als Prozessdaten bei den Prüforganisationen jährlich millionenfach anfallen und die Erwartung besteht, durch ihre Nutzung die Qualität der statistischen Information über Fahrleistungen insgesamt verbessern zu können.

Vor diesem Hintergrund kann zusammenfassend Folgendes festgestellt werden:

- Eine auf HU-Daten basierende laufende Fahrleistungsstatistik ist kein Ersatz für spezielle Fahrleistungserhebungen vom Typ der FLE 2002. Solche, in mehrjährigem Abstand durchzuführende, speziellen Fahrleistungserhebungen sollten ein Grundbaustein eines modernen verkehrstatistischen Informationssystems sein und bleiben.
- Eine auf HU-Daten basierende laufende Fahrleistungsstatistik kann gleichwohl aber eine Bereicherung für die amtliche Verkehrsstatistik darstellen.
- Obschon sich aus HU-Daten zum Tachostand nicht durch klassische Hochrechnung, sondern nur unter bestimmten - empirisch aber überprüfbaren (und hier auch überprüften!) -

Annahmen Schätzwerte für die kalenderjahrbezogene Fahrleistung von Kraftfahrzeugen generiert lassen, können die solchermaßen entstandenen Fahrleistungskennzahlen die Qualität der statistischen Information über die Inländerfahrleistung verbessern.

Der Beitrag einer auf HU-Daten basierenden laufenden Fahrleistungsstatistik zur Verbesserung der Informationslage über die Inländerfahrleistung ist wie folgt zu bewerten:

- Obschon sich in einer Hochrechnung von Tachostandsdaten zwangsläufig die gesamte Historie der Nutzung der betreffenden Stichprobenfahrzeuge widerspiegelt, können die Hochrechnungsergebnisse unter bestimmten Bedingungen als *hinreichend genaue Indikatoren für Volumen und Struktur der Kfz-Fahrleistungen eines aktuellen Kalenderjahres* angesehen werden.
- Die wesentliche Einschränkung, die hier zu machen ist, besteht darin, dass Strukturbrüche<sup>50</sup> in der zeitlichen Entwicklung der kalenderjahrbezogenen Gesamtfahrleistung, die aus *Veränderungen der mittleren Fahrleistung angemeldeter Fahrzeuge* am „aktuellen Rand“ resultieren, mit HU-Daten *nicht* diagnostiziert werden können. Wenn hier von der Veränderung der mittleren Fahrleistung angemeldeter Fahrzeuge gesprochen wird, so ist damit im Übrigen nicht das Kollektiv aller Fahrzeuge gemeint; gemeint sind vielmehr *Veränderungen der mittleren Fahrleistung innerhalb der einzelnen Teilgruppen von Fahrzeugen* (Hochrechnungsgruppen abgegrenzt durch die Merkmale Fahrzeugart, Fahrzeugalter, Antriebsart, Motorleistung/Ladepazität).
- Veränderungen der Gesamtfahrleistung am aktuellen Rand sind mit dem hier entwickelten Schätzverfahren nur dann quantifizierbar, soweit sie aus *Veränderungen im Fahrzeugbestand insgesamt* (Zahl und Strukturmerkmale der Fahrzeuge)) resultieren oder aus *Veränderungen in der Anmeldequote der Fahrzeuge* (Anmeldetage im Verhältnis zu allen Fahrzeugtagen eines Jahres).
- Da Umfang und Struktur des Fahrzeugbestands und ebenso die Anmeldequote als exogene Daten in das Hochrechnungsverfahren einfließen, spiegeln sich zeitliche Veränderungen dieser beiden Größen im Schätzwert für die kalenderjahrbezogene Fahrleistung unmittelbar wider. Dies bedeutet z.B., dass ein Fahrleistungsrückgang bei Nutzfahrzeugen, der aus konjunkturbedingt längeren Zeiten vorübergehender Stilllegung resultiert, vom hier entwickelten Schätzverfahren sehr wohl hinreichend genau wiedergegeben wird.

---

<sup>50</sup> Von einem „Strukturbruch“ kann im vorliegenden Zusammenhang gesprochen werden, wenn in einem linearen Regressionsmodell  $Y = \beta X + \varepsilon$  für die kalenderjahrbezogene Gesamtfahrleistung  $Y$  als Funktion des jahresdurchschnittlichen Fahrzeugbestands  $X$  der Parameter  $\beta$ , der in diesem Modell als mittlere aktuelle Jahresfahrleistung interpretiert werden kann, im Zeitverlauf nicht konstant ist, sondern sich nach oben oder unten (deutlich) ändert.

- Keinen hinreichend genauen Niederschlag im Schätzwert finden hingegen Veränderungen in der *gruppenspezifischen* mittleren Fahrleistung angemeldeter Fahrzeuge (z.B. Monatsfahrleistung von älteren Pkw mit Dieselmotor und einer Motorleistung bis 44 KW), wie sie z.B. als Reaktion auf drastisch gestiegene Kraftstoffpreise, sinkende Realeinkommen oder – noch extremer – als Folge von allgemeinen Fahrverboten im Gefolge einer Energiekrise vorkommen können.
- Für eine verlässliche Diagnose derartiger Veränderungen der mittleren Fahrleistung kommt letztlich nur ein *repräsentatives Fahrleistungspanel* in Betracht, in welches dann aber auch alle relevanten Fahrzeugarten einbezogen werden müssten.

Betrachtet man eine auf HU-Daten basierende Fahrleistungsstatistik als einen weiteren Baustein in einem statistischen Informationssystem, so könnte dieser Statistik eine relativ klar abgrenzbare Rolle zukommen:

- Die Ergebnisse einer neuen, auf HU-Daten basierenden Fahrleistungsstatistik sind zur Veröffentlichung in einer entsprechenden eigenen Reihe geeignet. Dort sollte allerdings klar auf die oben genannten Grenzen (aber auch auf das Potenzial) dieser Statistik hingewiesen werden. Unter diesen Voraussetzungen könnte die neue Fahrleistungsstatistik eigenständig für verkehrsplanerische und verkehrspolitische Zwecke genutzt werden.
- Die neue Fahrleistungsstatistik kann bislang nicht verfügbare Eingangs- bzw. Kontrollgrößen für das DIW-Berechnungsmodell zur Bestimmung von Fahrleistungen liefern.
- Neben Kennzahlen zum Niveau der jahresbezogenen Fahrleistung der verschiedenen Fahrzeugarten sind es vor allem Aufgliederungen der Fahrleistung nach technischen Fahrzeugmerkmalen, die mit den bisherigen Eingangsinformationen Kraftstoffabsatz, MiD-Daten und Daten des Mobilitätspanels nicht möglich sind.
- Eine auf HU-Daten basierende Fahrleistungsstatistik sollte ein eigenständiges System darstellen und im laufenden Betrieb nicht mit anderen laufenden Statistiken (z.B. Mobilitätspanel) „vermischt“ oder „angereichert“ werden. Naheliegender wäre es, die derzeitige DIW-Fahrleistungsrechnung zu einem Projekt „Metaanalyse Fahrleistungen“ zu erweitern, in dessen Rahmen die verschiedenen Informationsquellen zur Fahrleistung von Kraftfahrzeugen von einem Expertenteam bewertet und zusammengeführt werden.

Mit dem entwickelten Verfahren und der zugehörigen Software steht ein funktionsfähiges System zur Verfügung, das beim Kraftfahrt-Bundesamt in den laufenden Betrieb übergeführt werden kann. Hier ist aber klar zu sagen, dass nur eine „informierte“ Anwendung und Pflege des Systems in Betracht kommt. Die Produktion von Fahrleistungskennzahlen quasi „auf Knopfdruck“ ist eine illusorische Vorstellung. Zur Systempflege durch das beim KBA zu schaffende Team gehört auch die

periodische Überprüfung der im Zusammenhang mit dem Konzept der „neutralen“ Fahrleistung in das Modell eingeführten Korrekturfaktoren für die Fahrleistung vor dem ersten regulären HU-Termin; da diese Faktoren allesamt aus der Fahrleistungserhebung 2002 stammen, müsste hierfür allerdings eine Neuauflage dieser Erhebung vorliegen.

Methodenentwicklung ist wie jede wissenschaftliche Forschung niemals abgeschlossen. Insofern kann das hier vorgestellte Verfahren ganz sicher im einen oder anderen Punkt verbessert werden. Die *substantielle* Verbesserung einer auf Prozessdaten aus dem Kfz-Prüfwesen aufgebauten Fahrleistungsstatistik wird aber nicht durch zusätzliche hochrechnungsmethodische Feinheiten gelingen, sondern nur durch Veränderungen in den HU-Daten selbst: Wenn diese Daten den Tachostand nicht nur für die aktuelle sondern auch für die vorausgegangenen Hauptuntersuchungen (soweit durchgeführt) enthielten, läge im Sinne einer „verbundenen Stichprobe“ ein Datenmaterial vor, mit dem man dem Ziel einer sensitiven Veränderungsmessung einen entscheidenden Schritt näher kommen könnte.

## 12 Quellenverzeichnis

- BAST, Hrsg.(1978): *Fahrleistungen im Straßenverkehr - Datenbedarf, Erhebungsmethoden und Realisierungsempfehlungen zu Fahr- und Verkehrsleistungsstatistiken*. Autoren: Baur, R., Bihn, W.R., Brühning, E., Eichberg, J., Grunau, B., Otto, H., Rieke, H. und Zimmermann, G..Projektgruppenberichte der BAST, Bereich Straßenverkehrstechnik, Köln, Dezember 1978
- Hautzinger, H., Heidemann, D., Krämer, B. und Tassaux-Becker, B. (1994): *Fahrleistung und Unfallrisiko von Kraftfahrzeugen – Schlussbericht zur Fahrleistungserhebung 1990*. Berichte der BAST, Heft M 30
- Hautzinger, H. u.a. (2004): *Analyse von Änderungen des Mobilitätsverhaltens – insbesondere der Pkw-Fahrleistung – als Reaktion auf geänderte Kraftstoffpreise*. Schlussbericht zum BMVBW-Forschungsprojekt Nr. 96.0756/2002/. IVT: Heilbronn/Mannheim
- Hautzinger, H., Pfeiffer, M. und Schmidt, J. (2010): *Entwicklung eines methodischen Rahmenkonzepts für Verhaltensbeobachtung im fließenden Verkehr*. Schlussbericht zum Forschungsbericht FE 82.0342/2008 der BAST. IVT: Heilbronn/Mannheim
- Hautzinger, H., Stock, W., Mayer, K. Schmidt, J., Heidemann, D. (2005): *Fahrleistungserhebung 2002 - Inländerfahrleistung*. Berichte der BAST, Heft V 120
- Hautzinger, H., Stock, W. und Schmidt, J. (2005): *Fahrleistungserhebung 2002 - Inlandsfahrleistung und Unfallrisiko*. Berichte der BAST, Heft V 121
- Kalinowska, Dominika, Jutta Kloas, Hartmut Kuhfeld, Uwe Kunert (2005): *Aktualisierung und Weiterentwicklung der Berechnungsmodelle für die Fahrleistungen von Kraftfahrzeugen und für das Aufkommen und für die Verkehrsleistung im Personenverkehr (MIV)*. Gutachten des DIW Berlin im Auftrag des Bundesministeriums für Verkehr, Bau- und Wohnungswesen. Berlin, 2005.
- Kraftfahrt-Bundesamt: Statistische Mitteilungen, Reihe 7: Fahrzeuguntersuchungen. Flensburg.
- Kraftfahrt-Bundesamt: Statistische Mitteilungen, Fahrzeuguntersuchungen (FU). Hauptuntersuchungen sowie Einzelbegutachtungen und Änderungsabnahmen Jahr 2006. FU 1, Flensburg.
- Törkel, B. (1985): Künftige Anforderungen an die amtliche Statistik aus der Sicht der Verkehrspolitik. In: DVWG (Hrsg.), *Aktuelle Probleme und neue Methoden der Verkehrsstatistik*, Zweiter Workshop über Verkehrsstatistik - Leitung Prof. Dr. H. Hautzinger, DVWG-Schriftenreihe Band B 81, S. 64-90
- von der Lippe, P.M. (1990): *Wirtschaftsstatistik*. 4. Auflage, Gustav Fischer Verlag, Stuttgart