

Validität und Identifikation von Fahrzeugen und Fahrzeugeigenschaften in Kundenbefragungen

Jan Hendrik Schreier

ADAC e. V., Otto-Lilienthal-Straße 2, 86899 Landsberg am Lech, jan.schreier@tzll.adac.de

Prof. Dr. Niels Biethahn

Institut für Automobil Forschung (im RIF e.V.) und Professor für Unternehmenssteuerung und Projekt Manager für das Thema Automotive Management, BITS Business and Information Technology School gGmbH, Reiterweg 26b, 58535 Iserlohn, Niels.Biethahn@bits-iserlohn.de

1	Einführung und Problemstellung	236
2	Antwortvalidität bei Fragen zur Fahrzeugidentifikation	237
3	Auswirkung fehlerhafter und fehlender Antworten.....	240
4	Möglichkeiten, Grenzen und Risiken der Fahrzeugidentifikation.....	245
5	Fazit und Ausblick	246
6	Literaturverzeichnis	247

Abstract:

Für die Marktforschung der Automobilindustrie ist bei Zufriedenheitsbefragungen die eindeutige Identifikation einer Fahrzeugkonfiguration (z. B. Motorgröße und -art) von hoher Bedeutung für die Interpretation der Umfrageergebnisse. In diesem Paper wird untersucht, welche Kenntnisse Befragte zur Fahrzeugkonfiguration haben, wie valide die Aussagen von Befragten dazu sind und welche Möglichkeiten bestehen, um die Validität der Daten zu erhöhen.

JEL Classification: M31 (Marketing), O39 (Technological Change; Research and Development – Other)

Keywords: Market Research, Automotive, Recall Error, Car Identification

1 Einführung und Problemstellung

Die Entwicklung eines neuen Fahrzeugs, bzw. die Überarbeitung eines etablierten Modells, ist ein teures, mehrjähriges Verfahren (Gusig, Kruse, 2010: S. 29), dessen Erfolg erst nach Abschluss der Entwicklung anhand von Verkaufszahlen gemessen werden kann (Hagstotz, Schmitt-Hagstotz, 2003: S. 78). Um das Risiko kostspieliger Fehlentwicklungen zu reduzieren, werten die Automobilhersteller u. a. Zufriedenheitsstudien über aktuell im Markt befindliche fremde und eigene Modelle aus, um daraus Hinweise für die zukünftige Fahrzeugentwicklung abzuleiten (Diez, 2009: S. 73).

Von der Automobilmarktforschung wird bei Befragungen deshalb eine möglichst eindeutige und korrekte Fahrzeugidentifikation erwartet, damit aus den erhobenen Daten die richtigen Handlungen abgeleitet werden können. So macht es z. B. einen Unterschied, ob sich die Fahrer eines Pkw mit einer bestimmten Motorvariante über den Verbrauch beschweren oder ob es eine generelle Unzufriedenheit über alle Varianten eines Modells gibt.

Die Marktforschung ist daran interessiert, das zu beurteilende Fahrzeug möglichst detailliert und fehlerfrei zu identifizieren. Doch wie gut wissen Autofahrer über ihr Auto Bescheid? Während Modell und Marke den meisten Fahrzeughaltern noch geläufig sind, dürfte dies bei Leistung und Größe des Motors schon bei weniger Personen der Fall sein. Insbesondere, da einige Hersteller hier mit irreführenden Angaben arbeiten. So gibt der Hersteller BMW z. B. beim „BMW 335i“ den Hubraum mit 2.979 ccm an, obwohl die Bezeichnung nach der herstellereigenen Namenskonvention einen Wert von 3.500 ccm erwarten ließe (Horn, 2006). Diese Untersuchung trägt zu diesem Themenkomplex bei, indem sie Antworten auf die folgenden Fragen liefert:

1. Wie verlässlich werden welche Fahrzeuginformationen von Autofahrern in einer Befragung wiedergegeben?
2. Welche Auswirkungen können fehlerhafte Daten haben?
3. Welche Möglichkeiten und Grenzen gibt es, die Antwortqualität, insbesondere durch Hinzunahme der Daten aus der Zulassungsbescheinigung, zu verbessern?

Nicht diskutiert werden in dieser Untersuchung Fragestellungen zur Repräsentativität der Befragten im Verhältnis zur Grundgesamtheit. Siehe hierzu bspw. (Hahn, Jerusalem, 2003: S. 162–165), (Baur, Florian, 2009: S. 110–118), (Kutsch, 2007: S. 1–7), (Grunwald, Hempelmann, 2012: S. 49), (Thunig, 2007: S. 90–96). Die genannten Quellen führen als Hauptkritikpunkt die niedrige Quote von Personen mit Internetanschluss an. Die aktuelle Ausgabe des (N)onliner-Atlas zeigt jedoch, dass dieses Argument immer mehr an Bedeutung verliert (Initiative D21, 2012: S. 5).

Im zweiten Kapitel werden die Ergebnisse einer Befragung zur Fahrzeuginformation dargestellt, bei der ein Teil der Personen die Zulassungsbescheinigung zu Hilfe genommen und der andere Teil auf diese Hilfe verzichtet hat. Darauf aufbauend werden die sich ergebenden Probleme in Kapitel drei beleuchtet und ein kurzer Überblick über relevante Literatur zu diesem Problem gegeben.

Im vierten Kapitel werden Verfahren diskutiert, die zur Verbesserung der Antwortqualität beitragen können, sowie Grenzen und Risiken dieser Verfahren kurz skizziert. Im abschließenden fünften Kapitel wird weiterer Forschungsbedarf aufgezeigt.

2 Antwortvalidität bei Fragen zur Fahrzeugidentifikation

2.1 Einleitung

Kuß und Eißend definieren die Validität von Befragungen als gegeben, wenn der zu ermittelnde Sachverhalt von den erhobenen Daten tatsächlich wiedergegeben wird (2010: S. 30–31). Mit einer Marktforschungsstudie verfolgt der ADAC das Ziel, eine bestandsrepräsentative Erhebung über die Zufriedenheit mit Fahrzeugen zwischen einem halben und 3,5 Jahren zu erhalten. Dazu wurden 2012 erstmalig über 20.000 Fahrer mittels Onlinefragebogen befragt. Der Fragebogen hatte eine Ausfülldauer von durchschnittlich 20 Minuten und wurde über sogenannte Online-Panels¹ verteilt. Im ersten Teil des Fragebogens waren die Personen aufgefordert, ihr aktuelles Fahrzeug zu identifizieren. Hierzu mussten die Befragten u. a. folgende Informationen zu ihrem Fahrzeug preisgeben (in Klammern Beispielwerte):

- Marke (z. B. Audi)
- Modell (z. B. A4)
- Kraftstoffart (z. B. Diesel)
- Hubraum (z. B. 2,0 Liter)
- Motorleistung (z. B. 137 PS, wahlweise auch als kW-Angabe)
- Karosserieform (z. B. Limousine)

Im Pretest des Fragebogens wurde deutlich, dass sich nicht alle Personen an diese Daten korrekt erinnern können. Deshalb wurde der Fragebogen so erweitert, dass eine Fahrzeugidentifikation auch mittels Eingabe der in der Zulassungsbescheinigung stehenden Kombination von Hauptschlüsselnummer (HSN) und Typschlüsselnummer (TSN) möglich ist. Dies ist möglich, da in Deutschland bei Serienfahrzeugen die oben angegebenen Werte aus der HSN/TSN-Kombination abgeleitet werden können (Kraftfahrt-Bundesamt, 2012c: S. 3).

Die Angabe der eindeutigen HSN/TSN-Kombination in einer Befragung verlangt von den Befragten jedoch, dass Sie ihre Zulassungsbescheinigung zur Verfügung haben. Um die Abbruchquote der Befragung nicht zu erhöhen, wurde die Eingabe der HSN/TSN-Kombination in der Befragung nicht erzwungen, sondern als Erleichterung im Rahmen der Fahrzeugidentifikation angeboten. Alternativ konnten die Teilnehmer die Antworten mittels Einfachauswahl aus einer Liste auswählen bzw. im Fall der PS/kW-Angabe die entsprechende Zahl in ein Textfeld eingeben. Hierbei wurde lediglich das zu wählende Modell auf Basis der vorher ausgewählten Marke eingeschränkt. Die übrigen Werte wurden bei der freien Eingabe nicht eingeschränkt und es wurde in allen Fragen die Option „Weiß nicht“ angeboten. Im Fall von Marke und Modell führte diese Option jedoch zu einem Ausschluss von der Befragung.

¹ Unter einem Online-Panel versteht man einen „Pool von Personen, die sich bereit erklärt haben, an Online-Befragungen teilzunehmen“ (Arbeitskreis Deutscher Markt- & Sozialforschungsinstitute e. V., 2012).

Bei Eingabe einer HSN/TSN-Kombination wurde mittels Datenbankabfrage das entsprechende Fahrzeug identifiziert, so dass die Personen ihre Fahrzeugdaten nicht manuell eingeben mussten. Für die vorliegende Untersuchung wurde einer ausgewählten Teilmenge von Probanden mitgeteilt, dass die HSN/TSN-Kombination nicht gefunden wurde und die Daten zur Fahrzeugidentifikation manuell einzugeben seien.

Durch die Speicherung der vermeintlich nicht gefundenen HSN/TSN-Kombinationen war es möglich, die Werte der manuellen Eingaben und die sich aus HSN/TSN-Kombinationen ableitbaren Werte zu vergleichen. Hierdurch kann abgeschätzt werden, wie valide die Angaben einer manuellen Eingabe aller Einzeldaten ohne vorliegende Zulassungsbescheinigung sind.

Die vollständigen Interviews teilten sich dabei wie folgt auf:

	Anzahl	%
a) Es erfolgte keine Eingabe von HSN/TSN durch die Probanden.	13.595	67,9 %
b) Eine HSN/TSN-Kombination wurde eingegeben, war aber entweder falsch oder wurde als „nicht bekannt“ zurückgegeben. Die Personen mussten ihr Fahrzeug manuell identifizieren.	2.805	14,0 %
c) Eine HSN/TSN-Kombination wurde eingegeben und akzeptiert.	3.631	18,1 %
Summe	20.031	100 %

Table 1: Anteil der Befragten, die HSN/TSN-Kombinationen angegeben haben

2.2 Abgleich bei einer vorhandenen HSN/TSN-Kombination

In diesem Kapitel werden die 2.805 Fälle untersucht, bei denen sowohl eine HSN/TSN-Kombination eingegeben wurde, aber auch eine manuelle Eingabe stattfand. Von diesen 2.805 Fällen waren 1.826 Fälle gültige HSN/TSN-Kombinationen, wovon wiederum 53 Fälle ausgeschlossen wurden, da es sich hierbei nach KBA-Definition um „Leichte Nutzfahrzeuge“ handelte.

Für die 1.773 verbleibenden Fälle wurden die manuell eingegebenen Werte mit den Werten verglichen, die sich aus HSN/TSN-Kombination ableiten lassen. Für Marke, Modell, Hubraum, Motorleistung und Kraftstoffart konnten folgende Werte für korrekte Antworten ermittelt werden:

Variable	Korrekte Antworten
Marke	98,0 %
Modell	89,1 %
Motorleistung	78,3 %
Hubraum	85,2 %
Kraftstoffart	93,6 %

Table 2: Prozentsatz korrekter Antworten je Variable

Bei dem Fehleranteil der Modelle ist zu berücksichtigen, dass diese Antwort einen Folgefehler zur Marke enthält. Da die Antwortmöglichkeiten der Frage zum Modell dynamisch mit der Antwort zur Frage nach der Marke verknüpft ist, kann bei falscher Markenwahl das richtige Modell nicht ausgewählt werden.

Bei der Variablen Hubraum erfolgte der Abgleich mittels eines unscharfen Vergleichs, da die Befragten bei der Frage zum Hubraum aus einer Liste in 0,2 Liter-Schritten die entsprechende Größe auswählen mussten, in der KBA-Datenbank jedoch genaue Angaben enthalten sind. Dies führte in einigen Fällen dazu, dass z. B. ein Audi A4 2.0 (Liter) vom Befragten als solcher klassifiziert wurde, obwohl das Fahrzeug in Wirklichkeit nur über 1,968 Liter Hubraum verfügt und in die nächst kleinere Klasse sortiert wurde. Daher wurde bei dem Abgleich eine Angabe auch dann als korrekt bewertet, wenn der manuell eingegebene Wert und der sich aus der HSN/TSN-Kombination ergebende Wert in benachbarten Klassen lagen. Betrachtet man nur die Fälle mit größerer Abweichung, ist im Durchschnitt eine Überschätzung um 0,2 Liter zu messen.

Bei der Motorleistung wurden alle Werte auf kW umgerechnet und Abweichungen von weniger als 10 kW als korrekt eingestuft. Die durchschnittliche Abweichung von diesem Wert ist für alle falschen Antworten mit 0,15 kW sehr nah bei 0.

Bei der Karosserieform wurde auf die Angabe einer Prozentzahl korrekter Antworten verzichtet, da die weitere Analyse ergab, dass die Einordnung in den Daten des KBA wenig verlässlich ist, um die Karosserieform des Fahrzeugs zu identifizieren. Ein prominentes Beispiel hierfür ist der VW Golf VII, der in den Daten des KBA immer als „Kombilimousine“ identifiziert wird, von den meisten Fahrern aber als „Schrägheck“-Fahrzeug eingeordnet wird und nur dann auch als „Kombi“ bezeichnet wird, wenn es sich um die längere Variante mit größerem Kofferraum handelt.

Betrachtet man die Kombination aller Variablen je Befragten, ergibt sich die nachfolgend abgebildete Verteilung:

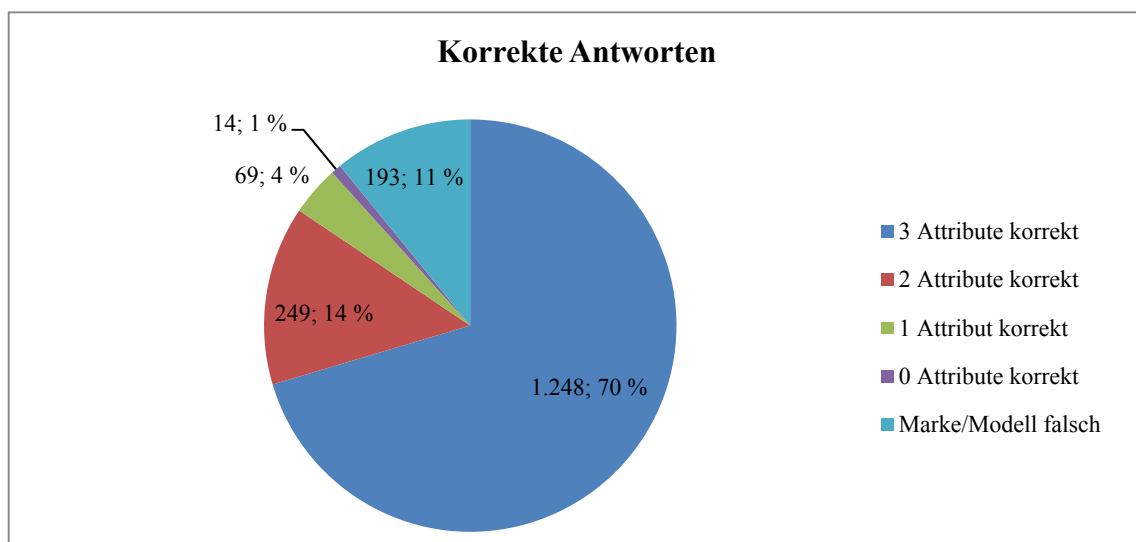


Figure 1: Häufigkeitsverteilung der korrekten Antworten

In 1.248 (70 %) der 1.773 Fälle hat der oder die Befragte die Attribute Marke, Modell, Hubraum, Motorleistung und Kraftstoffart korrekt identifiziert. In 249 Fällen (14 %) war ein Attribut falsch. In allen Fällen ist davon auszugehen, dass zumindest ein Teil der Personen die Daten aus der vorliegenden Zulassungsbescheinigung entnommen hat, da eine korrekte Eingabe einer HSN/TSN-Kombination ohne vorliegende Zulassungsbescheinigung extrem unwahrscheinlich ist und im Fragebogen Beispielabbildungen enthalten waren, die darauf hin-

weisen, an welcher Stelle in der Zulassungsbescheinigung die jeweilige Information zu finden ist.

2.3 Überprüfung der Validität bei Befragten ohne HSN/TSN-Eingabe

Um zu ermitteln, wie verlässlich diese Information ist, wurde für Fahrer der Marken Dacia und Audi untersucht, wie valide die Antworten der Personen sind, die auf die Eingabe einer HSN/TSN-Kombination verzichtet haben (Gruppe a in Table 1). Dazu wurde geprüft, ob die manuell eingegebenen Informationen in der jeweiligen Kombination vorkommen können.

Hierzu wurden die Angaben der Befragten mit den KBA-Daten abgeglichen. Dabei zeigte sich, dass die Antworten der Dacia-Fahrer etwas weniger valide waren als die Antworten der Audi-Fahrer. Eine mögliche Interpretation wäre ein scheinbar geringeres Interesse am Thema Auto bei den Dacia-Fahrern. Der Abgleich beschränkte sich hier auf die Attribute Hubraum und Motorleistung, da für alle Modelle sowohl Benzin- als auch Dieselmotoren zur Verfügung standen und der Abgleich deshalb keine Auffälligkeit zeigen konnte. Der Prozentsatz valider Antworten je Marke und Variable verteilt sich dabei wie folgt:

Variable	Audi	Dacia
Motorleistung	68 %	64 %
Hubraum	80 %	66 %

Table 3: Valide Antworten bei Befragung ohne HSN/TSN-Kombination

Die Werte entsprechen in der Tendenz denjenigen aus dem vorherigen Kapitel. Der größte Teil (jeweils über 80 %) der Personen, die falsche Werte angegeben hat, hat hier allerdings keinen unplausiblen Wert angegeben, sondern direkt die Kategorie „Weiß nicht“ gewählt. Im Vergleich zu den Personen, die ihre Zulassungsbescheinigung vorliegen hatten, ist der Anteil an Antworten in der Kategorie „Weiß nicht“ mehr als doppelt so hoch.

Betrachtet man die vollständig korrekt identifizierten Fahrzeuge, erreicht man für alle Audi A3 einen Wert von 55 %, für Dacia Logan 38 % und für den Dacia Sandero einen Wert von 48 %.

2.4 Zwischenfazit

Wie nicht anders zu erwarten, ist die Antwortqualität bei vorliegender Zulassungsbescheinigung deutlich höher. Mit Zulassungsbescheinigung liegt die Fehlerquote je nach Attribut zwischen 10 und 21 %. Ohne Zulassungsbescheinigung liegt die Fehlerquote je Attribut zwischen 20 und 34 %.

Die Untersuchung zeigt außerdem die Grenzen des Einsatzes von HSN/TSN-Kombinationen insbesondere zur Festlegung der Karosserieform. Hier erscheint eine Auswahl durch den Befragten mittels skizzierten Formen verlässlicher als die Datenbasis des KBA.

3 Auswirkung fehlerhafter und fehlender Antworten

In der Literatur findet man zum Thema Rückerinnerung aktuell keine Untersuchungen im Bereich Automotive. Schwerpunkte der Forschung bilden z. B. Untersuchungen zur Abschät-

zung von Aussagen durch Befragte in der Konsumgüterforschung (z. B. Hoderlein, Winter, 2007) oder Fragen zum Abstimmverhalten in der Wahlforschung (z. B. Schoen, 2009). Für die Bundestagswahl 2005 liegt die Rückerinnerung der gewählten Partei nach 2 bis 3 Jahren bei etwas über 60 % (Schoen, 2009: S. 266), also in der gleichen Größenordnung wie der Anteil korrekter Antworten in dieser Untersuchung.

Bound u. a. benennen als Auswirkungen falscher Antworten, dass 1) reale Effekte versteckt bleiben, 2) Zusammenhänge zu existieren scheinen, die es nicht gibt und 3) Koeffizienten das falsche Vorzeichen erhalten können (2001: S. 3708–3709). Die Effekte 2) und 3) nach Bound u. a. können nur durch falsche Antworten entstehen, die symmetrisch, also nicht zufällig sind, während 1) auch bzw. insbesondere durch eine „Weiß nicht“-Antwortoption begünstigt wird.

Zur Abschätzung der Fehlerrate sind entweder starke Annahmen über die Fehlerursachen oder Validierungsdaten nötig. Letztere erlauben es, weniger starke Annahmen zu treffen. Bei den zusätzlichen Datenquellen gibt es vier Optionen:

Der Abgleich kann

- a) mit einer weiteren Datenquelle je Befragtem,
- b) durch Messung der Varianz in den Antworten bei Befragungen der gleichen Personen zu unterschiedlichen Zeitpunkten,
- c) mit einer später erhobenen Datenquelle je Befragtem oder
- d) über den Abgleich auf Makroebene (z. B. offiziellen Statistiken)

erfolgen (Bound u. a., 2001: S. 3740–3742).

Im Falle der untersuchten Befragung kann der Rückgriff auf die Daten, die sich aus HSN/TSN-Kombination ergeben als Variante a) verstanden werden. Ein Abgleich mit den KBA-Bestandsdaten im Sinne der Variante d) wäre ebenfalls möglich, die Abweichungen wären dann jedoch nicht eindeutig einem Fehler zuzuweisen, sondern z. B. auch durch das Sample der Befragung zu erklären.

Für den Fall der Motorleistung erfolgt die fehlerhafte Beantwortung nahezu normalverteilt. Auswertungen, in die die Motorleistung als erklärende Variable einfließt, werden daher eine größere Streuung feststellen, Prognosen werden dadurch ungenauer.

Für die Variable Hubraum ist der Einfluss kritischer, da hier bei den Falschantworten eine Überschätzung erkennbar ist. Bei den 15 % der Falscheingaben mit verfügbaren Daten aus HSN/TSN-Kombination wurde z. B. der Durchschnitt über alle Fahrzeuge um etwa 0,12 Liter zu hoch berichtet. Der von Horn (Horn, 2006) beschriebene Umstand, dass BMW hier mit seiner Modellbezeichnung den Kunden verwirrt, konnte in der Erhebung festgestellt werden. Entgegen der Annahme, dass BMW-Fahrer besonders autoaffin sind, war hier die Fehlerquote mit 24,3 % deutlich höher als der Durchschnitt über alle Marken. Für die Marke Mercedes, die eine ähnliche Namenspolitik wie BMW verfolgt (Horn, 2006), lag der Effekt mit 17,5 % nur etwas über dem Durchschnitt.

Zur Überprüfung der These, dass die Fehlangaben die Unsicherheit der Ergebnisse erhöhen, wird getestet, ob die Fehler in einem signifikanten Zusammenhang zur Zufriedenheit stehen. Es wird die Hypothese H_0 aufgestellt:

Es besteht kein empirisch messbarer Zusammenhang zwischen der angegebenen Hubraumgröße, der Korrektheit der Hubraumangabe und der Zufriedenheit mit „Antrieb/Motor insgesamt“.

Ziel ist es, die Hypothese H_0 zu verwerfen und dadurch zu zeigen, dass die Korrektheit der Angabe einen messbaren Einfluss auf die Zufriedenheit hat. Die Korrektheit für die Hubraumangabe je Fall wurde in einer neuen Variablen als Über- bzw. Unterschätzung und korrekte Werte kodiert:

$$\text{Korrektheit}_{\text{Hubraumangabe}} := \begin{cases} -1 = \text{Hubraum unterschätzt} \\ 0 = \text{Hubraum korrekt} \\ 1 = \text{Hubraum überschätzt} \end{cases}$$

Formel 1: Faktor Korrektheit der Hubraumangabe

Die aufgestellte Regressionsgleichung lautet:

$$\text{Gesamtzufriedenheit}_{\text{Motor/Antrieb}} = b_0 + b_1 * \text{Hubraumangabe} + b_2 * \text{Korrektheit}_{\text{Hubraumangabe}}$$

Formel 2: 1. Regressionsgleichung Gesamtzufriedenheit mit Motor/Antrieb

Der F-Test für dieses Regressionsmodell ergibt einen Wert von 91,293 (siehe Table 4). Bei 2 und 1749 Freiheitsgraden bedeutet dies einen hoch signifikanten Zusammenhang ($p < 0,001$) für das Gesamtmodell. Die Nullhypothese kann verworfen und von einem Zusammenhang zwischen der angegebenen Hubraumgröße, der Korrektheit der Hubraumangabe und der Zufriedenheit mit „Antrieb/Motor insgesamt“ ausgegangen werden. Die Erklärung der Streuung für dieses Modell ist mit 9,5 % der Gesamtstreuung relativ niedrig. Bezogen auf den Mittelwert zur Gesamtzufriedenheit mit Antrieb/Motor von 2,40 beträgt der Standardfehler des Schätzers 33 % ($= 0,801/2,40$), was nicht als gut bewertet werden kann (Backhaus u. a., 2006: S. 73). Es gibt also weitere Faktoren, die zur Zufriedenheit mit dem Motor/Antrieb beitragen, die hier aber nicht weiter von Interesse sind.

Modellzusammenfassung									
Modell	R	R ²	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderungsstatistiken				
					Änderung in R-Quadrat	Änderung in F	df1	df2	Sig. Änderung in F
1	0,307 ^a	0,095	0,093	0,80100	0,095	91,293	2	1749	0,000

a. Einflussvariablen: (Konstante), Korrektheit_{Hubraumangabe}, Hubraumangabe

Table 4: Modellzusammenfassung der Regressionsgleichung zur Motor/Antriebs-Gesamtzufriedenheit mit Hubraumangabe und Korrektheit der Hubraumangabe

Die Hubraumangabe hat hochsignifikanten Einfluss auf die Zufriedenheit mit dem Motor/Antrieb. Die standardisierten Koeffizienten zeigen, dass die Hubraumangabe von den beiden unabhängigen Variablen den größeren Einfluss auf die Zufriedenheit mit dem Motor/Antrieb insgesamt hat, da sie mit 0,322 als absolutem Wert größer ist als die Korrektheitsangabe mit 0,048 (siehe Table 5).

Der Faktor zur Korrektheit der Hubraumangabe verfehlt bei einem t-Test die Signifikanzgrenze ($p < 0,05$) mit 0,052 nur knapp, die Hubraumangabe selbst ist im t-Test hochsignifikant. Der Faktor für die Hubraumangabe ist negativ, da die Zufriedenheit mit dem Motor/Antrieb auf einer Skala von 1 = „Ausgezeichnet“ bis 5 = „Schlecht“ bewertet wurde und die Hub-

raumangaben von 1 = „unter 1 Liter“ und 0,2 Liter-Schritten um den Wert 1 erhöht wurden. Das bedeutet, dass größerer Hubraum für steigende Zufriedenheit mit dem Antrieb sorgt, was plausibel erscheint. Das 95 %-Konfidenzintervall für den Koeffizienten zur Korrektheit der Hubraumangabe liegt zwischen -0,001 und 0,216 und erlaubt daher keine inhaltliche Interpretation, da ein Vorzeichenwechsel im Konfidenzintervall vorkommt (Backhaus u. a., 2006: S. 77–78). In diesem Fall kann die Hypothese H_0 nicht verworfen werden, weil der Faktor zur Korrektheit nicht signifikant ist.

Koeffizienten ^a						
Modell	Nicht standardisierte Koeffizienten			Standardisierte Koeffizienten	T	Sig.
	Regressionskoeffizient B	Standardfehler	Beta			
1	(Konstante)	2,849	0,039		73,358	0,000
	Hubraumangabe	-0,090	0,007	-0,322	-13,148	0,000
	Korrektheit _{Hubraumangabe}	0,108	0,055	0,048	1,946	0,052

a. Abhängige Variable: Frage F27 - Motor / Antrieb - Wie beurteilen Sie den Motor/Antrieb des Fahrzeugs insgesamt?

Table 5: Koeffizienten der Regressionsgleichung mit Über- bzw. Unterschätzung der Hubraumangabe

In einem zweiten Modell wird zusätzlich die Interaktion zwischen der Hubraumangabe und deren Korrektheit untersucht und die Regressionsgleichung erweitert:

$$\text{Gesamtzufriedenheit}_{\text{Motor / Antrieb}} = b_0 + b_1 * \text{Hubraumangabe} + b_2 * \text{Korrektheit}_{\text{Hubraumangabe}} + b_3 * \text{Interaktion}_{\text{Korrektheit, Hubraumangabe}}$$

Formel 3: 2. Regressionsgleichung Gesamtzufriedenheit mit Motor/Antrieb

Der Interaktionsfaktor ist definiert als:

$$\text{Interaktion}_{\text{Korrektheit, Hubraumang.}} := \text{Hubraumangabe} + \text{Korrektheit}_{\text{Hubraumangabe}}$$

Formel 4: Interaktionsfaktor Hubraumangaben und Korrektheit der Hubraumangabe

Die Regressionsgleichung für das Gesamtmodell ist wiederum hochsignifikant. Da durch die Hinzunahme weiterer Variablen ein Anstieg des R^2 immer gegeben ist, ist hier ein Vergleich der Werte für das korrigierte R-Quadrat nötig (Backhaus u. a., 2006: S. 68). Dieser Wert steigt von 0,093 für die erste Regressionsgleichung auf 0,102 für die zweite, um den Interaktionseffekt erweiterte, Regressionsgleichung (siehe Table 6). Die Berücksichtigung der Interaktion verbessert die Aussagekraft des Modells. Der Standardfehler des Schätzers sinkt nur leicht, die Abweichung zum Mittelwert beträgt weiterhin 33 %.

Modellzusammenfassung									
Modell	R	R ²	Korrigiertes R-Quadrat	Standardfehler des Schätzers	Änderungsstatistiken				
					Änderung in R-Quadrat	Änderung in F	df1	df2	Sig. Änderung in F
1	0,322 ^a	0,104	0,102	0,79719	0,104	67,360	3	1748	0,000

a. Einflussvariablen: (Konstante), Korrektheit_{Hubraumangabe}, Hubraumangabe, Interaktion_{Korrektheit, Hubraumang.}

Table 6: Modellzusammenfassung der Regressionsgleichung zur Motor/Antrieb-Gesamtzufriedenheit mit Hubraumangabe, Korrektheit der Hubraumangabe und Interaktion zwischen Hubraumangabe und deren Korrektheit

Der Faktor zur Korrektheit der Hubraumangabe wird signifikant ($p = 0,027$; siehe Table 7), der Interaktionsfaktor und die angegebene Hubraumgröße sind hochsignifikant. Inhaltlich gilt weiterhin, dass größerer Hubraum zu größerer Zufriedenheit mit dem Motor/Antrieb führt. Der Koeffizient zur Korrektheit ist jetzt auch im 95 %-Konfidenzintervall eindeutig negativ (-0,385; -0,023). Da ein kleiner Wert für die Gesamtzufriedenheit Motor/Antrieb höhere Zufriedenheit ausdrückt und der Faktor zur Korrektheit negativ bei Unterschätzung bzw. positiv bei Überschätzung des Hubraums ist, besagt dieser Faktor, dass ein Fahrer umso glücklicher ist, je größer er seinen Hubraum einschätzt. Da die Richtung der Abhängigkeit nicht eindeutig ist, kann es auch sein, dass der Hubraum umso größer geschätzt wird, je zufriedener ein Autofahrer mit seinem Motor/Antrieb insgesamt ist. Der Faktor zur Interaktion weist in die entgegengesetzte Richtung wie die beiden anderen Faktoren, reduziert also den Effekt einer Über- bzw. Unterschätzung etwas. Die standardisierten Koeffizienten zeigen, dass die Hubraumangabe den größten Einfluss hat (-0,366), gefolgt vom Interaktionsfaktor (0,186). Die Korrektheitsangabe hat den geringsten Einfluss (-0,090).

Koeffizienten ^a					
Modell	Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Sig.
	Regressionskoeffizient B	Standardfehler	Beta		
1 (Konstante)	2,892	0,040		72,360	0,000
Hubraumangabe	-0,102	0,007	-0,366	-13,803	0,000
Korrektheit _{Hubraumangabe}	-0,204	0,092	-0,090	-2,213	0,027
Interaktion _{Korrektheit, Hubraumang.}	0,047	0,011	0,186	4,213	0,000

a. Abhängige Variable: Frage F27 - Motor / Antrieb - Wie beurteilen Sie den Motor/Antrieb des Fahrzeugs insgesamt?

Table 7: Koeffizienten der Regressionsgleichung mit Hubraumangabe, Korrektheit der Hubraumangabe und Interaktion zwischen Hubraumangabe und deren Korrektheit

Da sowohl das Modell als auch alle Faktoren in den Tests signifikante Ergebnisse aufweisen, kann mit diesem Modell die These H_0 verworfen werden. Die aufgestellte These wird akzeptiert: Es besteht ein positiver Zusammenhang zwischen der Überschätzung der Hubraumgröße und der Zufriedenheit mit dem Motor/Antrieb.

Die vorgestellte Analyse sowie die ermittelten Fehlerquoten können als konservativ gelten. Es ist davon auszugehen, dass die meisten Personen in dieser Untersuchung ihre Zulassungsbescheinigung vorliegen hatten, da eine korrekte HSN/TSN-Eingabe ohne Hilfe der Zulassungsbescheinigung äußerst unwahrscheinlich ist. Die Signifikanz, die in der vorliegenden Untersuchung teilweise durch die hohe Fallzahl begründet ist, wird daher bei niedrigerer Fallzahl, aber höherer Fehlerzahl bei Antworten ohne Verwendung der Zulassungsbescheinigung, weiter Bestand haben.

Um das Problem von Falscheingaben für Befragungen zu verringern, können die Daten schon während der Eingabe validiert werden (Treiblmaier, 2011: S. 8). Eine Validierung durch Einschränkung der Antwortauswahl scheint kritisch, da Menschen dazu tendieren, sich bei (personenbezogenen) Fragen in der Mitte einzusortieren (Hoderlein, Winter, 2007: S. 40–43). In den in dieser Untersuchung untersuchten und ähnlichen Fällen scheint es daher sinnvoller, die

Antwortmöglichkeiten nicht einzuschränken, bei unplausiblen Werten jedoch erneut nachzufragen und die Befragten ggf. auf die Antwortoption „Weiß nicht“ hinzuweisen.

Neben den bisher beschriebenen Problemen weist Treiblmaier auch auf das Problem der Glaubwürdigkeit hin (2011: S. 5–6). Dies ist insbesondere von praktischer Relevanz, da eine geringe Glaubwürdigkeit die Akzeptanz der Ergebnisse und davon abhängige Änderungsimpulse deutlich reduziert.

4 Möglichkeiten, Grenzen und Risiken der Fahrzeugidentifikation

4.1 Verfahren zur Fahrzeugidentifikation

Neben den beiden bereits genannten Verfahren der direkten Befragung oder der Ableitung von Informationen aus HSN/TSN-Kombinationen gibt es als dritte Variante zur Fahrzeugidentifikation den Rückgriff auf die Fahrzeugidentifikationsnummer (FIN). Über die FIN kann jedes Fahrzeug weltweit eindeutig identifiziert werden (DIN, 1985: S. 3). In den folgenden Kapiteln werden die drei Verfahren, inklusive möglicher Vorteile und Risiken, kurz beschrieben.

4.2 Direkte Abfrage der Daten von Befragten

Die direkte Abfrage der Daten ist technisch die einfachste Variante, jedoch auch die fehleranfälligste. Zur Verringerung der Fehlerrate kann man die Befragten bitten, die Zulassungsbescheinigung zu Rate zu ziehen. Für die eindeutige Identifikation eines Facelifts oder eines neuen Modells (z. B. die Unterscheidung zwischen Golf VI und Golf VII) bietet sich der Einsatz von Fotos der jeweiligen Fahrzeuge an. Dies ist jedoch mit erheblichem Aufwand verbunden und zur Verlässlichkeit der Zuordnung liegen noch keine Informationen vor. Eine Einschränkung der Antwortkategorien im Falle der Motorleistung oder -größe erscheint kritisch, da auf diese Weise Falschantworten erzeugt werden könnten.

4.3 HSN/TSN-gestützte Befragung

Der Einsatz der HSN/TSN-Kombination zur Ableitung von Fahrzeuginformationen ist, wie in Kapitel 0 gezeigt, nicht frei von Problemen. Neben der genannten Schwäche bei der Karosserieform ist auch die Unterscheidung von Facelifts teilweise nicht möglich, da ein Facelift nicht zwingend eine neue HSN/TSN-Kombination erhält². Im Gegensatz zur freien Eingabe durch den Befragten hat die HSN/TSN-Kombination jedoch den Vorteil, dass die Daten zur Kraftstoffart, Motorgröße und -leistung relativ zuverlässig identifiziert werden können.

Ein weiteres Problem bei der Identifikation stellt die TSN „000“ dar. Diese wird für Fahrzeuge vergeben, die mittels Einzeltypgenehmigung zugelassen werden (z. B. stark modifizierte Fahrzeuge oder Kleinstserienfahrzeuge) oder in Deutschland nicht typisiert sind (z. B. Importfahrzeuge, die so auf dem deutschen Markt nicht verfügbar sind) (Kraftfahrt-Bundesamt, 2012b: S. 1). Für die TSN „000“ ist also lediglich die Marke ableitbar.

² Zum Beispiel ist die HSN/TSN-Kombination 0603/AMK sowohl dem Golf V als auch dem Golf VI zugeordnet. In der KBA-Liste wird er nur als „Golf“ geführt (Kraftfahrt-Bundesamt, 2012c: S. 154).

Die These, dass Befragte mittels kurzer Suchmaschinenabfrage eine gültige HSN/TSN-Kombination ermitteln, konnte für die vorliegende Untersuchung verworfen werden. Möchte man die Rate der Personen, die HSN/TSN eingeben, erhöhen und incentiviert die Eingabe, wäre diese These jedoch erneut zu prüfen.

4.4 Einsatz der Fahrzeugidentifikationsnummer

Die FIN besteht aus der Welt-Herstellerbezeichnung (WMI – World Manufacturer Identification), dem fahrzeugbeschreibenden Teil (VDS – Vehicle Description Section) und dem fahrzeugunterscheidenden Teil (VIS – Vehicle Identification Section). Die VIS ist eine fortlaufende Nummer und stellt sicher, dass die FIN eindeutig ist. Über WMI und VDS lassen sich, je nach Hersteller, unterschiedliche Werte eindeutig bestimmen (Schwacke, 2012: S. 2794–2795). Die Anzahl der hinterlegten, öffentlich verfügbaren Attribute schwankt jedoch zwischen den Herstellern. Während z. B. Audi nur das Modell, Modelljahr, Herstellungsjahr und -werk angibt (Schwacke, 2012: S. 2718–2720), gibt BMW zusätzlich noch die Motorgröße sowie Motor- und Karosserietyp an (Schwacke, 2012: S. 2721).

Eine generelle Verwendung der FIN zur Identifikation von Fahrzeugen mehrerer Hersteller ist zwar teilweise möglich, aber mit sehr hohem Aufwand verbunden. Das Modell- bzw. Baujahr ist auch nicht bei allen Herstellern lesbar in der FIN kodiert, da es innerhalb der Vorgaben nur freiwillig ist, diesen Wert zu speichern. Das Gesetz verlangt lediglich, dass ein Fahrzeug über 30 Jahre vom Hersteller eindeutig identifiziert werden kann (Europäische Union, 1976).

Hauptproblem beim Einsatz der FIN ist, dass die vollständigen Daten zum Fahrzeug nur beim jeweiligen Hersteller selbst vorliegen. Die FIN verfügt in deutschen Zulassungsbescheinigungen über eine Prüfziffer, so dass die Korrektheit schon bei der Eingabe überprüft werden kann (Kraftfahrt-Bundesamt, 2012a: S. 1–2).

4.5 Zwischenfazit

Zur vollständigen Fahrzeugidentifikation, inklusive aller Sonderausstattungen, ist die FIN theoretisch das beste Verfahren. Es setzt jedoch voraus, dass alle wesentlichen Hersteller Zugriff auf ihre internen FIN-Listen gewähren. Dies wird nur in den seltensten Fällen praktisch umsetzbar sein. Da aber selbst im Optimalfall nicht alle befragten Personen bereit sein dürften, ihre FIN einzugeben, wäre auch in diesem Fall der Einsatz eines oder beider vorher genannten Verfahren nötig.

5 Fazit und Ausblick

In dieser Untersuchung wurde das Wissen über Attribute des eigenen Pkw bei Fahrern junger Fahrzeuge empirisch untersucht und Werte für die Validität der Antworten ermittelt. Dabei wurde gezeigt, dass die Validität und Vollständigkeit der Antworten bei vorliegender Zulassungsbescheinigung deutlich höher ist.

Die ermittelten Werte wurden mit Werten anderer wissenschaftlicher Untersuchungen verglichen und praktische Implikationen auf Basis der eigenen Werte ermittelt. Der sich daraus er-

gebende Handlungsdruck wurde aufgenommen und weitere Optionen und deren Risiken zur Fahrzeugidentifikation skizziert.

Das Thema Fahrzeugidentifikation ist von hoher praktischer Relevanz, da eine niedrige Antwortvalidität die Akzeptanz der Ergebnisse jeder Studie verringert (Deshpande, Zaltman, 1982: S. 25) und die Chance auf Handlungsimpulse – das eigentliche Ziel der Marktforschung – schmälert. Daher gibt es weiteren Forschungsbedarf wie Befragte motiviert werden können, ihre Zulassungsbescheinigung beim Ausfüllen des Fragebogens zu Rate zu ziehen.

Neben der Identifikation der in dieser Untersuchung besprochenen Aspekte, gewinnen Fahrerassistenzsysteme aktuell stark an Bedeutung. Das Wissen über die jeweils im Fahrzeug installierten Systeme ist jedoch gering (Deutscher Verkehrssicherheitsrat, 2012: S. 1). Aus diesem Grund ist hierzu eine Untersuchung zur Antwortvalidität wünschenswert.

6 Literaturverzeichnis

Arbeitskreis Deutscher Markt- & Sozialforschungsinstitute e. V. (2012): „Online-(Access-)Panels“. Abgerufen am 12.08.2012 von <https://www.adm-ev.de/index.php?id=onlinepanels>.

Backhaus, Klaus; Erichson, Bernd; Plinke, Wulff; u. a. (2006): *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. 11. Aufl. Berlin [u.a.]. — ISBN: 978-3-540-85044-1 3-540-85044-9

Baur, Nina; Florian, Michael J. (2009): „Stichprobenprobleme bei Online-Umfragen“. In: *Sozialforschung im Internet: Methodologie und Praxis der Online-Befragung*. Wiesbaden, S. 109–128.

Bound, John; Brown, Charles; Mathiowetz, Nancy (2001): „Measurement Error in Survey Data“. In: Heckmann, J. J.; Leamer, E. (Hrsg.) *Handbook of Econometrics*. Amsterdam.

Deshpande, Rohit; Zaltman, Gerald (1982): „Factors Affecting the Use of Market Research Information: A Path Analysis“. In: *Journal of Marketing Research*. 19 (1), S. 14–31.

Deutscher Verkehrssicherheitsrat (2012): „Fahrerassistenzsysteme helfen älteren Verkehrsteilnehmern“. Abgerufen am 15.07.2012 von <http://www.dvr.de/presse/informationen/1462.htm>.

Diez, Willi (2009): *Automobil-Marketing. Navigationssystem für neue Absatzstrategien*. 1. Auflage. München.

DIN (1985): „DIN ISO 3779 - Fahrzeug-Identifizierungsnummer (VIN)“.

Europäische Union (1976): „Richtlinie 76/114/EWG“.

Grunwald, Guido; Hempelmann, Bernd (2012): *Angewandte Marktforschung. Eine praxisorientierte Einführung*. München.

Gusig, Lars-Oliver; Kruse, Arne (2010): *Fahrzeugentwicklung im Automobilbau. Aktuelle Werkzeuge für den Praxiseinsatz*. München.

Hagstotz, Werner; Schmitt-Hagstotz, Karin (2003): „Automobil-Marktforschung: Status und neue Herausforderungen“. In: *Planung & Analyse*. Jg. 2003 (5), S. 78–84.

Hahn, A.; Jerusalem, M. (2003): „Reliabilität und Validität in der On-line-Forschung“. In: Online-Marktforschung: Theoretische Grundlagen und praktische Erfahrungen. Wiesbaden, S. 161–186.

Hoderlein, Stefan; Winter, Joachim (2007): „Recall Errors in Surveys“. Abgerufen am 08.01.2012 von <http://www.econ.yale.edu/seminars/ecm/ecm07/Hoderlein-071003.pdf>.

Horn, Jan (2006): „Turbolente Zeiten“. [autobild.de](http://www.autobild.de). Abgerufen am 22.08.2012 von <http://www.autobild.de/artikel/fahrbericht-bmw-335i-coupe-56497.html>.

Initiative D21 (2012): „(N)onliner Atlas 2012“. Abgerufen am 03.07.2012 von <http://www.initiaved21.de/wp-content/uploads/2012/06/NONLINER-Atlas-2012-Basiszahlen-f%C3%BCr-Deutschland.pdf>.

Kraftfahrt-Bundesamt (2012a): „Berechnung der Prüfwert zur FIN nach dem Modulo-11-Verfahren“. Abgerufen am 15.08.2012 von http://www.kba.de/nr_239644/DE/ZentraleRegister/ZFZR/Info__behoerden/anlage__2__Berechnung__Pruefwert__FIN__Modulo__11__Verfahren__pdf,templateId=raw,property=publicationFile.pdf/anlage__2__Berechnung__Pruefwert__FIN__Modulo__11__Verfahren__pdf.pdf.

Kraftfahrt-Bundesamt (2012b): „Emailantwort des Kraftfahrt-Bundesamts“.

Kraftfahrt-Bundesamt (2012c): „Verzeichnis der Hersteller und Typen der für die Personenbeförderung ausgelegten und gebauten Kraftfahrzeuge mit mindestens vier Rädern (Klasse M)“. Abgerufen am 16.08.2012 von http://www.kba.de/nr_125156/DE/Fahrzeugtechnik/Dienstleistungen__VeroeffentlichungenTGV/Veroeffentlichungen/veroeffentlichungen__inhalt__SV4.html.

Kuß, Alfred; Eisend, Martin (2010): Marktforschung; Grundlagen der Datenerhebung und Datenanalyse. 3., überarb. und erw. Aufl. Wiesbaden.

Kutsch, Horst B. (2007): Repräsentativität in der Online-Marktforschung. Lösungsansätze zur Reduktion von Verzerrungen bei Befragungen im Internet. 1. Aufl. Lohmar [u.a.].

Schoen, Harald (2009): „Wenn ich mich recht erinnere... Zur Validität von Rückerinnerungsfragen“. In: Vom Interview zur Analyse. Methodische Aspekte der Einstellungs- und Wahlforschung. Baden-Baden, S. 259–279.

Schwacke (2012): Schwacke Liste - Typenerkennung 2012-13 Band II. o.V.

Thunig, Christian (2007): „Probanden dringend gesucht!“. In: Absatzwirtschaft. 50 (Heft 3), S. 90–97.

Treiblmaier, Horst (2011): „Datenqualität und Validität bei Online-Befragungen“. In: *der markt*. 50 (1), S. 3–18, DOI: 10.1007/s12642-010-0030-y.