

## ZI-Praxis-Panel und amtliche Kostenstrukturerhebung: ein Methodenvergleich

Peter Michael von der Lippe

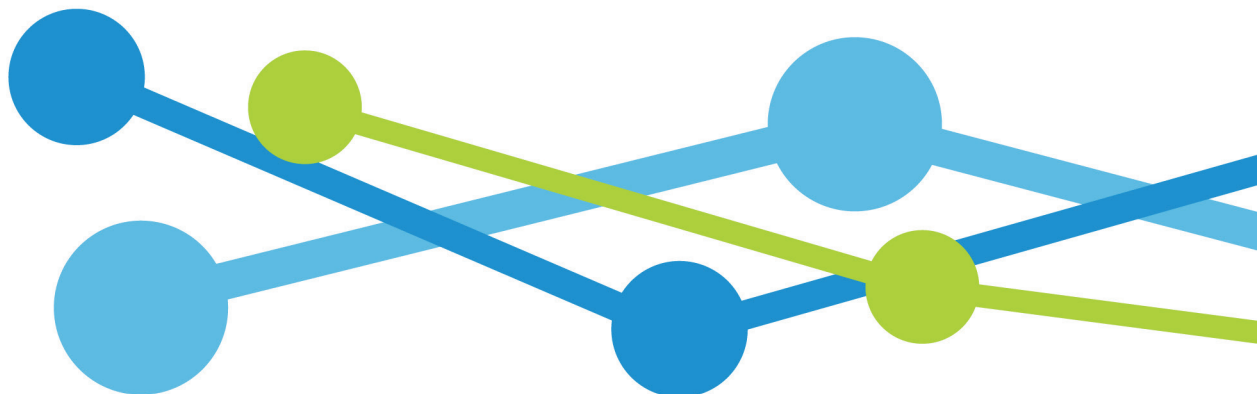


[www.zi.de](http://www.zi.de)

**Zentralinstitut für die  
kassenärztliche Versorgung  
in der Bundesrepublik Deutschland**

Herbert-Lewin-Platz 3  
10623 Berlin

Tel. +49 30 4005 2450  
Fax +49 30 4005 2490  
[zi@zi.de](mailto:zi@zi.de)



# ZI-Praxis-Panel und amtliche Kostenstrukturerhebung: ein Methodenvergleich

Prof. Dr. Peter Michael von der Lippe  
Professor (Emeritus) für Statistik, Universität Duisburg-Essen

## Vorbemerkung

Die folgende Darstellung vergleicht das ZI-Praxis-Panel (kurz ZiPP) mit der amtlichen Kostenstrukturerhebung (kurz KSE) bei Arzt-, Zahnarzt- und Tierarztpraxen des Statistischen Bundesamts [abgekürzt StBA oder Destatis]. Die das ZiPP betreffenden Informationen stammen aus Veröffentlichungen und aus persönlichen Gesprächen mit Mitarbeitern des ZI.<sup>1</sup> Die Angaben zur KSE sind weitgehend dem Qualitätsbericht des StBA entnommen.<sup>2</sup>

---

1 ZI-Praxis-Panel, Jahresbericht 2010 – Wirtschaftliche Situation und Rahmenbedingungen in der vertragsärztlichen Versorgung der Jahre 2006 bis 2008, 1. Jahrgang, Berlin, März 2012. Bei der Zusammenstellung des Textes wurde mir ganz erheblich von Mitarbeitern des Zentralinstituts für die kassenärztliche Versorgung in der Bundesrepublik Deutschland (ZI) geholfen, die auch die Richtigkeit meiner hier gemachten Angaben noch einmal überprüft haben. Ich danke in diesem Zusammenhang insbesondere Herrn Markus Leibner. In einigen Abschnitten, insbesondere Abschn. 1.3, 2.2, 2.4 und 4.2 konnte ich auf frühere Ausarbeitungen, die ich im Auftrag des ZI machte, zurückgreifen.

2 Qualitätsbericht Kostenstrukturerhebung bei Arzt- und Zahnarzt- und Tierarztpraxen 2007, August 2009. Detailinformationen zur KSE sind auch entnommen aus Fachserie 2, Reihe 1.6.1, Unternehmen und Arbeitsstätten Kostenstruktur bei Arzt- und Zahnarztpraxen, Praxen von psychologischen Psychotherapeuten sowie Tierarztpraxen 2007, erschienen am 5. August 2009 Artikelnummer: 20201610790. Grundlage für die Darstellung waren auch der Aufsatz Taubmann D. u. K. Harrer, Kostenstrukturen bei Arzt- und Zahnarztpraxen sowie Praxen von psychologischen Psychotherapeutinnen und -therapeuten 2007, in: Wirtschaft und Statistik 4/2010, S. 333 – 343 und Gespräche mit den beiden Verfasserinnen. Hinsichtlich des Stichprobenfehlers hat es auch einen Schriftwechsel mit Herrn Dipl. Mathematiker Kai Lorentz im Statistischen Bundesamt gegeben.

## Inhaltsverzeichnis

<b>1.</b>	<b>Grundlagen der beiden Befragungen.....</b>	<b>7</b>
1.1	Auftrag und Zielsetzung der Erhebungen .....	7
1.2	Erhebung und Nutzung weiterer Datenquellen beim ZiPP.....	8
1.3	Panel-Konzept und Standardisierung der Einnahmen des ZiPP .....	10
1.4	Zusammenfassung der Besonderheiten des ZiPP.....	12
<b>2.</b>	<b>Daten zu den Stichproben (Auswahlverfahren, Erhebung, Hochrechnung) .....</b>	<b>13</b>
2.1	Bereichsabgrenzung und Auswahlgesamtheit .....	13
2.2	Stichprobenplanung und Durchführung der Stichprobenziehung .....	15
2.3	Fehlersystematik, Antwortausfälle (nicht-stichprobenbedingte Fehler).....	17
2.4	Hochrechnung.....	20
2.5	Stichprobenfehler, Genauigkeit und „Repräsentativität“ .....	22
<b>3.</b>	<b>Einzelheiten zum Merkmalskatalog .....</b>	<b>24</b>
3.1	Grobübersicht zu den Merkmalskatalogen der beiden Befragungen.....	24
3.2	Struktur der Praxen, Versorgungsleistungen .....	26
3.3	Einnahmen, Ausgaben und Vermögen im Detail .....	27
<b>4.</b>	<b>Besondere methodische Aspekte.....</b>	<b>30</b>
4.1	Bemerkungen zur Panelbefragung .....	30
4.2	Standardisierung der Einnahmen als Besonderheit des ZiPP .....	32
<b>5.</b>	<b>Zusammenfassung .....</b>	<b>38</b>
	<b>Tabellenanhang .....</b>	<b>40</b>

## Verzeichnis der Übersichten

	<b>Titel</b>	<b>Seite</b>
1	Allgemeine Angaben zu Art und Zweck der Erhebung	7
2	Angaben zur Erhebungstechnik und Datenverknüpfung	9
3	Abgrenzung des Erhebungsbereichs und Auswahlgesamtheit	15
4	Übersicht zu den Merkmalskatalogen der beiden Befragungen	25
5	Praxistyp, Unternehmensdemographie und Versorgungsdaten	26
6	Personelle, räumliche und technische Ausstattung der Praxis	27
7	Details zu Fragen nach Einnahmen und Aufwendungen	28

## Verzeichnis der Tabellen

	<b>Titel</b>	<b>Seite</b>
1	Daten zum Auswahlrahmen (Vergleich KSE/ZiPP)	14
2	Anzahl der befragten Praxen und Rücklauf der Fragebogen (Vergleich KSE/ZiPP)	18
3	Geschätzte Anzahl der Antwortausfälle in der KSE	20
4	Hochgerechnete und noch nicht hochgerechnete Häufigkeiten im ZiPP	21
5	Relativer Stichprobenfehler in Prozent bei Kernindikatoren der Kostenstrukturerhebung	22
6	Relativer Stichprobenfehler in Prozent bei Praxisaufwendungen 2008 (ZiPP)	23
7	Relativer Stichprobenfehler mit Endlichkeitskorrektur in Prozent bei Praxisaufwendungen 2008	23
8	Impliziter Faktor $\lambda \cdot \kappa$ bei einigen standardisierten Jahresüberschüssen (ZiPP)	34
9	Standardisierungsfaktor $\lambda$ bei den befragten Arztpraxen (ZiPP)	35
10	Faktor $\alpha$ bei Arztpraxen ausgewählter Fachrichtungen (ZiPP)	35
A1	Daten zur Stichprobenziehung	40
A2	Praxen der Grundgesamtheit differenziert nach den 108 Schichten im ZiPP	41
A3	Differenzierung nach Arztgruppen in der KSE und im ZiPP	42
A4	Faktor $\alpha$ bei Arztpraxen ausgewählter Fachrichtungen (alle Praxen umfassend, auch in Spalte 1 reine GKV-Praxen und Praxen bei denen $\alpha = 1$ gesetzt wurde)	43

## Verzeichnis der Abbildungen

	<b>Titel</b>	<b>Seite</b>
1	Fehlersystematik bei statistischen Erhebungen	19
2	Häufigkeitsverteilung der $\alpha$ -Werte (ohne reine GKV-Praxen und Praxen bei denen $\alpha$ auf 1 gesetzt wurde)	36
3	Häufigkeitsverteilung der $\alpha$ -Werte (mit reinen GKV-Praxen und Praxen bei denen $\alpha$ auf 1 gesetzt wurde)	43

## Verzeichnis der Abkürzungen

	Titel
BAG	Berufsausübungsgemeinschaft (= Gemeinschaftspraxis)
BAR	Bundesarztregister
BSNR	Betriebsstättennummer
DMP	Disease Management Programme
EBM	Einheitlicher Bewertungsmaßstab
GKV	Gesetzliche Krankenversicherung
KBV	Kassenärztlichen Bundesvereinigung
KSE	Kostenstrukturerhebung
KV	Kassenärztliche Vereinigung(en)
LANR	Lebenslange Arztnummer
MVZ	Medizinisches Versorgungszentrum
NACE	Nomenclature statistique des activités économiques dans la Communauté Européenne*
PKV	Private Krankenversicherung
SGB	Sozialgesetzbuch
StBA	(oder Destatis) Statistisches Bundesamt
URS	Statistisches Unternehmensregister
VDX	Vereinheitlichter Datenaustausch in XML (Abrechnungsdaten der KVen und der KBV)
VGR	Volkswirtschaftliche Gesamtrechnung(en)
ZI	Zentralinstitut für die kassenärztliche Versorgung in der Bundesrepublik Deutschland
ZiPP	ZI-Praxis-Panel

\*) Statistische Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft, ihr entspricht der WZ 2003 (Klassifikation der Wirtschaftszweige – Ausgabe 2003) in Deutschland

## 1. Grundlagen der beiden Befragungen

In diesem Abschnitt werden die Erhebungen bezüglich ihrer allgemeinen Kennzeichen verglichen. Die folgenden Kapitel (insbesondere 2 und 3) enthalten dann detailliertere Angaben zur Stichprobenmethodik und zu den in den beiden Erhebungen jeweils erhobenen Merkmalen.

### 1.1 Auftrag und Zielsetzung der Erhebungen

Zu den allgemeinen Kennzeichen von Erhebungen gehören Name, Zweck und Ziel der Erhebung, Erhebungstermin und Periodizität, Erhebungsgesamtheit, -einheiten und -gegenstände (Merkmale) usw. Die hierfür wichtigsten Angaben finden sich in Übers. 1 und 2. Nach Darstellung der allgemeinen Kennzeichen des ZiPPs im Vergleich zur KSE werden in Abschnitt 1.4 die Besonderheiten des ZiPP noch einmal zusammengefasst und kommentiert.

#### Übersicht 1

Allgemeine Angaben zu Art und Zweck der Erhebung<sup>3</sup>

	KSE (StBA)	ZiPP (ZI)
Name der Erhebung	Kostenstrukturstatistik bei Arzt-, Zahnarzt- und Tierarztpraxen (EVAS-Nr.: 52571).	ZI-Praxis-Panel (ZiPP), eingeführt (erste Befragung) 2010.
Rechtsgrundlage	Gesetz über Kostenstrukturstatistik (KoStrukStatG) 7.9.2007 (BGBl. I S. 2246); amtliche Statistik mit Auskunftspflicht.	Beschluss des Kuratoriums des ZI am 27.9.2009, das ZI mit dem Panel zu beauftragen; Teilnahme ist freiwillig.
Zweck der Erhebung	Ergänzung von Statistiken, die primär das Ergebnis der Wirtschaftstätigkeit (Auftragseingang, Produktion, Umsatz) betrachten und für die VGR wichtig sind.	Unterstützung von Honorarverhandlungen, Sicherstellung d. vertragsärztlichen Versorgung (daher auch viele versorgungsrelevante Merkmale).
Erhebungsart	Wiederholte Querschnitte (schriftliche Befragung); erfragte Angaben beziehen sich nur auf das Erhebungsjahr (2007), geschichtete Stichprobe, ca. 5.300 Praxen davon 4.090 auswertbare Fragebögen.	Panelerhebung (schriftliche Befragung); alle Angaben differenziert für jedes der drei Jahre 2006, 2007 und 2008. Geschichtete Stichprobe mit 4.661 Teilnehmern im Jahr 2010 und einem angestrebten Umfang von ca. 7.000 Praxen
Periodizität	vierjährlich, zuletzt 2003, 2007, 2011	erstmalig 2010, wird jährlich wiederholt

Das ZiPP trägt einem 2009 durch neue gesetzliche Regelungen (§ 87 SGB V) ausgelösten Informationsbedarf nach Angaben zum Niveau und zur Entwicklung (auch hinsichtlich regionaler Besonderheiten) der Kosten der niedergelassenen Ärzteschaft Rechnung. Daneben besteht auch Informationsbedarf für das Standardbewertungssystem STABS, um für die Aktualisierung des einheitlichen Bewertungsmaßstab (EBM) Informationen über Kosten und Kostenarten je Fachgruppe zur Verfügung zu stellen. Der Zweck der amtlichen KSE ist demgegenüber sehr viel allgemeiner. Sie verfolgt insbesondere nicht das Ziel, eine vergleichbare Wiederholungsbefragung (Panelerhebung) zur wirtschaftlichen Lage der Praxen niedergelassener Ärzte zu sein und ihre Ergebnisse sind nicht gedacht als unmittelbar relevant für Vergütungsentscheidungen. Genau das ist aber bei den Feststellungen des ZiPP zur Kostenentwicklung

<sup>3</sup> Die Übersicht enthält, anders als eine ähnliche frühere synoptische Darstellung des ZI, keine Angaben zur früheren jährlichen „ZI - Kostenstrukturanalyse in Arztpraxen“. Das war eine Erhebung, die keine Panelerhebung war und die durch das ZiPP ersetzt wurde.

der Fall. So ist beispielsweise bei der Anpassung der bundeseinheitlichen Punktwerte als Orientierungswerte gemäß § 87 Absatz 2f SGB V) „die Entwicklung der für Arztpraxen relevanten Investitions- und Betriebskosten“ zu berücksichtigen.<sup>4</sup>

Das ZI prüfte diverse Datenbestände (z.B. KSE, DATEV, apoBank) hinsichtlich der Eignung der in

ihnen enthaltenen Informationen über die wirtschaftliche Situation von niedergelassenen Ärzten oder Psychotherapeuten<sup>5</sup> für die genannten Zwecke. Es stellte sich heraus, dass keine dieser Quellen laufend die gewünschten Informationen mit ausreichender Qualität bieten kann. Das rechtfertigt die Existenz einer speziellen nichtamtlichen Erhebung in Gestalt des ZiPP neben der amtlichen KSE.

## 1.2 Erhebung und Nutzung weiterer Datenquellen beim ZiPP

Als nichtamtliche Statistik ist die Teilnahme beim ZiPP freiwillig. Es war daher mit Nichtbeantwortung zu rechnen. Die KSE ist demgegenüber eine amtliche Statistik und es herrscht Auskunftspflicht der Erhebungseinheiten (Praxen).

Andererseits kann das ZiPP Datenquellen zur Erhebung und Validierung der Erhebungsergebnisse nutzen, die der KSE nicht zur Verfügung stehen. Beide Erhebungen sind geschichtete Stichproben.<sup>6</sup> Die ZiPP-Stichprobe wurde aus dem Bundesarztregister (BAR) der KBV gezogen, das dem Statistischen Bundesamt für die KSE nicht zur Verfügung stand. Auswahlgrundlage war dort das statistische Unternehmensregister (URS).

Das URS enthält Angaben zur wirtschaftlichen Tätigkeit und zur Größe (Zahl der sozialversicherungspflichtig Beschäftigten)<sup>7</sup> der erfassten Einheiten und es dient nicht nur als Auswahl-, sondern auch als Hochrechnungsrahmen sowie als Quelle für Strukturdaten (es enthält auch Angaben zur Aufnahme bzw. Einstellung sowie wirtschaftszweigsystematischen Zuordnung der Tätigkeit der Einheit). Beim Aufbau des URS verwendete Datenquellen sind vor allem Angaben der Arbeitsverwaltung (sozialversicherungspflichtige Beschäftigte).<sup>8</sup>

Die ZiPP-Befragung 2010 umfasste die Jahre 2006, 2007 und 2008. Dass die Jahre bei der Erhebung bereits etwas weiter zurück lagen, liegt daran, dass viele Angaben der steuerlichen Überschussrechnung zu entnehmen waren.

<sup>4</sup> Dies kann erfolgen in Form von Angaben zur Kostenentwicklung pro Praxis oder pro Praxisinhaber, die bereitzustellen ebenfalls Ziel des ZiPP ist.

<sup>5</sup> Bei Psychotherapeuten sind ärztliche und psychologische Psychotherapeuten zu unterscheiden. Letztere zählen zwar nicht zu den Ärzten, sind aber dennoch im BAR (Bundesarztregister) geführt. Sie werden im ZiPP wie auch bei der KSE befragt.

<sup>6</sup> Bei einer geschichteten Stichprobe wird die Grundgesamtheit vollständig in Schichten zerlegt und es wird aus jeder Schicht eine Zufallsauswahl getroffen.

<sup>7</sup> Das URS enthält kaum Umsatzdaten, da Ärzte üblicherweise nur selten umsatzsteuerpflichtig sind. Angaben der Finanzverwaltung (Umsatzsteuer) konnten also zum Aufbau des URS im Falle von Ärzten nur selten genutzt werden und als Indikator für die Praxisgröße und zur Bildung der Schichten konnte somit nicht der Umsatz, sondern nur die Zahl der sozialversicherungspflichtig Beschäftigten herangezogen werden.

<sup>8</sup> Praxen ohne sozialversicherungspflichtig Beschäftigte sind deshalb nicht in der Grundgesamtheit enthalten.



## Übersicht 2

Angaben zur Erhebungstechnik und Datenverknüpfung

	KSE (StBA)	ZIPP (ZI)
Art der Auswahl (geschichtete Stichprobe)	Auswahlsatz <sup>a)</sup> ca 5%, Schichtung nach Gebiet (alte/neue Bundesländer), Wirtschaftszweig <sup>b)</sup> (Facharztbezeichnung), Zahl der sozialversicherungspflichtigen Beschäftigten aus dem statistischen Unternehmensregister (URS)	Schichtung nach drei Regionen und 36 Fachgruppen [genauer: Abrechnungsgruppen] (108 = 3 · 36 Schichten). Auswahl aus d. Bundesarztregister (BAR) der KBV Stand 31.12.2008 (aktualisiert zum 31.10.2009)
Erhebungszeitraum	Berichtszeitraum war das Kalenderjahr 2007 bzw. das maximal 12 Monate dauernde Geschäftsjahr, das im Laufe des Kalenderjahres 2007 endete <sup>c)</sup>	Die erfragten Daten (u.a. der steuerl. Überschussrechnung) betreffen die Jahre 2006, 2007 und 2008. Daten beziehen sich jeweils auf ein Kalenderjahr
Erhebungseinheiten	Praxen (freie Berufe) in Bereichen gem. NACE, die 2007 aktiv waren und nicht zu den öffentlich-rechtlichen Einheiten zählten	Praxen und Ärzte der vertragsärztlichen Versorgung ("Kassenarztpraxen"), die vor dem 1.1.2008 bestanden bzw. tätig waren
Durchführung der Erhebung	zentrale Erhebung des Statistischen Bundesamts (StBA); Stichprobenziehung im April 2008; Fragebogen am 26.9.2008 durch StBA verschickt	Fragebogenversand April/Mai 2010; Zwischenschaltung eines externen Treuhänders, Zahlung einer Aufwandspauschale bei Teilnahme
Merkmale	Strukturdaten (grob), Einnahmen, Ausgaben, Beschäftigte, mehr Details vgl. Übers. 4; Erhebungsinhalte primär die Kosten (nur dort ähnlich wie ZiPP)	viele (versorgungsrelevante) Merkmale der Praxis, Einnahmen, Ausgaben, Beschäftigte, Praxisausstattung <sup>d)</sup> , Kapital; Testat des Steuerberaters
Regionale Gliederung	Deutschland und Differenzierung nach 1. Früheres Bundesgebiet, 2. Neue Länder und Berlin-Ost	Darstellung nach Regionstypen des BBSR <sup>e)</sup> <ul style="list-style-type: none"> <li>• Agglomerationsräume</li> <li>• verstädterte Räume und</li> <li>• ländliche Räume</li> </ul>
Verknüpfungsmöglichkeiten	Keine weiteren Primärerhebungen des StBA in den durch die KSE erfassten Bereichen, keine Bezüge zu anderen amtlichen Erhebungen	Kombination mit Abrechnungsdaten der KBV möglich (VDX-Daten); dadurch auch viele Plausibilisierungsmöglichkeiten und weitere Auswertungsmöglichkeiten

a) systematische Zufallsauswahl als Auswahltechnik

b) NACE als Wirtschaftszweigklassifikation

c) Die Angaben für das Merkmal „Tätige Personen“ sind stichtagsbezogen. Der Stichtag war der 30.9.2007.

d) weitgehend gleicher Fragebogen, aber in Teilen (z.B. Geräte) 16 unterschiedliche Fassungen je nach Fachrichtung

e) Kreistypen 1,5 (Agglomerationsräume), 4,7,9 (verstädterte Räume) und 2,3,6 und 8 (ländliche Räume)

Zu den Besonderheiten des ZIPP gehört, dass bei allen finanziellen Angaben ein Testat des Steuerberaters verlangt wurde. Weitere Besonderheiten<sup>9</sup> des ZiPP gegenüber der KSE neben der Einbeziehung des Steuerberaters (die auch die Zahlung einer Aufwandsentschädigung bei Teilnahme erforderlich machte) sind auch

9 Vgl. Abschn. 1.4 zu einer zusammenfassenden Darstellung der Besonderheiten des ZiPP.

1. die Zwischenschaltung eines externen Treuhänders zur Wahrung der Anonymität der Teilnehmer,
2. die kleinräumige (Gemeinden) Erfassung und Darstellung der Daten und
3. die Möglichkeit der Kombination der ZiPP Ergebnisse mit Abrechnungsdaten der KBV, den sog. VDX-Daten,<sup>10</sup> d.h. von Abrechnungsdaten der KVen für ärztliche Leistungen.

Die Abrechnungsdaten beinhalten Leistungs-, Honorar- und Strukturdaten der Praxen wie

- Fachgebiet der Praxis (Zuordnung nach dem Abrechnungsverhalten der Praxis)
- Organisationsform (Einzel- / Gemeinschaftspraxis / MVZ)
- Anzahl der Leistungserbringer (d.h. Anzahl der tätigen Inhaber und angestellten Ärzte)
- Honorareinnahmen der Praxis<sup>11</sup>

sowie patienten- oder fallbezogene Informationen wie

- Fallzahl (cum grano salis: „Anzahl der Krankenscheine“)
- EBM (Gebührenordnungs-)Positionen<sup>12</sup> und
- Alter und Geschlecht des Versicherten (Patienten).

### 1.3 Panel-Konzept und Standardisierung der Einnahmen des ZiPP

Im ZiPP wird versucht, in besonderem Maße die Vergleichbarkeit von Kosten- und Ertragsdaten von Praxen sicherzustellen. Vergleichbarkeit von statistischen Gesamtheiten (z.B. die Arztpraxen einer Fachgruppe in einer Region) verlangt, dass sich die Einheiten der Gesamtheit (d.h. die Praxen) möglichst nicht durch Unterschiedlichkeit hinsichtlich vergleichsstörender Merkmale unterscheiden.

Beim zeitlichen Vergleich der Kosten und Erträge von Praxen ist die Vergleichbarkeit sicher dann im besonderen Maße gewährleistet, wenn es sich im Zeitablauf stets um genau die gleichen Praxen handelt, die erhoben werden und wenn somit die ermittelten Durchschnittswerte für bestimmte Kosten oder Relationen (z.B. Reinerträge je Praxisinhaber) nicht durch signifikante Veränderungen der Gesamtheit (Hinzukommen und Ausscheiden von Praxen, gravierende Strukturveränderungen innerhalb der Praxen, z.B. Veränderungen im Personal oder im Leistungsspektrum) verfälscht sind. „Reine“ Vergleiche durchzuführen verlangt den Einfluss von Strukturveränderungen bzw. -unterschieden auszuschalten, sei es, indem man

- sich auf die Betrachtung von (Teil-) Gesamtheiten beschränkt, die im Zeitablauf tatsächlich keine Veränderung (hinsichtlich vergleichsstörender Umstände) erfahren, oder indem man
- rein rechnerisch den Einfluss von Strukturveränderungen bzw. -unterschieden auf die ermittelten Kenngrößen (z.B. auf die absoluten Gesamtumsätze oder eine Verhältniszahl, wie z.B. die Rentabilität) eliminiert (was man in der Statistik „Standardisierung“ nennt).<sup>13</sup>

Beide Punkte betreffen zwei weitere Besonderheiten des ZiPP gegenüber der KSE

1. das ZiPP ist als Panelerhebung konzipiert und

10 Der Ausdruck VDX leitet sich ab aus: „Vereinheitlichter Datenaustausch in XML“ und bezeichnet eine Datenbank mit Abrechnungsdaten der Vertragsärzte und -psychotherapeuten. „Vereinheitlicht“ bezieht sich auf die Angleichung von regionalen Modifikationen der Abrechnungsverfahren in einzelnen KVen. Die Möglichkeit, ZiPP und VDX zu verknüpfen, ist über die Betriebsstättennummer BSNR gegeben.

11 Gemeint ist damit das von der KV gezahlte Honorar, das sofern nicht an Selektivverträgen teilgenommen wurde im ZiPP der Position „B2.1 Einnahmen aus ambulanter und belegärztlicher GKV-Praxis“ entspricht.

12 EBM Positionen, sind Ziffern zur Identifizierung einer Leistung aus dem EBM-Katalog. In Verbindung mit den vorgegebenen Kalkulationszeiten im EBM können die – für die Standardisierung (vgl. Absch. 4.2) wichtige – Jahresleistungszeiten des Arztes plausibilisiert werden.

13 In diesem Sinne spricht man z.B. von einer standardisierten Sterberate. Die „rohe“ Sterberate (Gestorbene eines Jahres dividiert durch die Wohnbevölkerung) eines Landes A ist nicht vergleichbar mit der rohen Sterberate eines Landes B, weil z.B. die Altersstruktur in A ungünstiger sein kann (relativ mehr alte Menschen) als in B und allein deswegen die rohe Sterberate von A größer sein kann als die von B. Um das zu berücksichtigen werden standardisierte Sterberaten berechnet, d.h. es wird die Sterblichkeit in beiden Länder auf Basis der gleichen Altersstruktur miteinander verglichen.

2. im Zuge der Auswertung werden neben den erhobenen Einnahmen auch standardisierte Einnahmen berechnet und veröffentlicht.

Zu 1: Echte Längsschnittsbetrachtungen zur Darstellung von Entwicklungen frei von Strukturveränderungen sind sehr aufwändige Statistiken,<sup>14</sup> die somit auch entsprechend selten durchgeführt werden. Eine wesentliche Voraussetzung derartiger Erhebung ist das Vorhandensein von im Zeitablauf konstanten Identifikationsmerkmalen. Dies ist nötig, um festzustellen, dass es sich bei Daten einer Praxis in  $t$ ,  $t+1$ ,  $t+2$ , ... jeweils um Daten der gleichen Praxis für die Erhebungs- und für die Untersuchungseinheit handelt. Diese Voraussetzung ist im Falle des ZiPP gegeben, weil die Betriebsstättennummer (BSNR) und die lebenslange Arztnummer (LANR) eine entsprechende Identifikation durch Vergabe einer konstanten Nummer erlaubt.<sup>15</sup>

Zu 2: Die Nutzung des ZiPP zur Anpassung des Orientierungswertes und der EBM Bewertungsrelationen macht Umrechnungen der erhobenen Daten über Einnahmen (Umsätze) und Reinerträge erforderlich, weil nicht die tatsächlich beobachtete Arztpraxis, sondern eine fiktive Praxis als „Normpraxis“ bei entsprechenden Verhandlungen zugrunde gelegt wird. Die Bestimmung der Einnahmen einer solchen Praxis ist Gegenstand der „Standardisierung“, die im Falle des ZiPP, nicht aber bei der KSE vorgenommen wird. Die „Normpraxis“ in diesem Sinne ist gekennzeichnet durch

- einen Privatpatientenanteil in Höhe von  $p = 0$  (reine GKV-Praxis) und
- einer jährlichen Arbeitszeit je Arzt in Höhe von  $A^* = 140.148$  Minuten.

Zweck der Standardisierung ist es, den Einfluss eines tatsächlichen Privatpatientenanteils in Höhe von  $p > 0$  und einer tatsächlichen Arbeitszeit  $A \neq A^*$  auf die Einnahmen zu bestimmen. In diesem Sinne eliminiert die Standardisierung den Einfluss unterschiedlicher einnahmerelevanter Strukturen (z.B. hinsichtlich der Patientenstruktur nach dem Merkmal PKV/GKV und der Arbeitszeit) auf die Einnahmen und ermöglicht somit einen „bereinigten“ Vergleich.

Zwar dürfte es mit gesundheitspolitischen Zielsetzungen zu rechtfertigen sein, dass man sich bei der von der GKV zu leistenden Vergütung an einer entsprechenden Norm- oder Normalpraxis orientiert, die eine akzeptable Einnahme erzielen würde, auch dann, wenn sie nur „Kassenpatienten“ hätte, damit ist es aber noch keineswegs klar, wie die notwendig „fiktiven“ Einnahmen einer solchen Praxis zu bestimmen sind. Auf die hiermit verbundenen methodischen Probleme wird in Abschnitt 4.2 eingegangen.

Der strukturelle Einfluss unterschiedlicher Arbeitszeiten dürfte dabei weniger umstritten sein als der Einfluss der Patientenstruktur.

Weicht das Niveau der „standardisierten“ Einnahmen einer (fiktiven) reinen GKV-Praxis vom tatsächlich ermittelten Durchschnitt aus den „realen“ Praxen (die jeweils einem mehr oder weniger großen Anteil von Privatpatienten haben) erheblich ab, stellt sich natürlich die Frage, ob die angewandte Standardisierungsmethode korrekt ist.

Die im Rahmen des ZiPP angewandte Methode geht davon aus, dass eine Praxis umso höhere Einnahmen im Vergleich zur Normpraxis erzielt, je größer

- der Anteil  $p$  der Privatpatienten ist, und je mehr
- das Verhältnis  $\alpha$  der Einnahme je PKV-Patient ( $e_p$ ) im Verhältnis zur entsprechenden „Pro-Kopf“ Einnahme  $e_G$  je GKV-Patient ist also je weiter  $\alpha = e_p/e_G$  über 1 liegt.

Es erscheint uns für die Akzeptanz der Standardisierung, die naturgemäß kritisch betrachtet werden dürfte, besonders wichtig, dass

<sup>14</sup> Es gibt eine Reihe weiterer Untersuchungsziele, die dann auch eine andere Herangehensweise bei der Wiederholungsbefragung verlangen. Vgl. hierzu einige weiterführende Bemerkungen in Abschn. 4.2.

<sup>15</sup> Auf Basis der beiden Nummern wurde im Rahmen der Pseudonymisierung eine spezielle Nummer für die ZiPP Teilnehmer kreiert.

- die Berechnung nur auf rein definatorischen Zusammenhängen zwischen Einnahmen (pro Patient), dem Anteil  $p$  der Privatpatienten und der Relation  $\alpha = e_P/e_G$  beruht und dass somit keine Modellannahmen über das Verhalten von Ärzten oder Patienten erforderlich sind, und dass
- sich die für die Standardisierung nach der Versichertenstruktur maßgeblichen Größen  $p$  und  $\alpha$  allein aus im ZiPP erhobenen Daten errechnen lassen.

#### 1.4 Zusammenfassung der Besonderheiten des ZiPP

Die folgenden Punkte dürften als wesentliche Kennzeichen des ZiPP im Vergleich zur amtlichen Kostenstrukturerhebung (KSE) besonders bemerkenswert sein:

- Das ZiPP enthält im Unterschied zur KSE nur Praxen der kassenärztlichen Versorgung (*Kassenarztpraxen*), es ist
- eine *Panelbefragung* (bei einer Befragung werden Daten zu jeweils drei Jahre erfragt, und es wird eine wiederholte Befragung der gleichen Einheiten angestrebt; die dazu notwendigen Voraussetzungen einer Identifizierbarkeit gleicher Einheiten im Zeitablauf über Arzt- und Betriebsstättennummern sowie eigens im Zuge der Erhebung vergebenen Nummern sind erfüllt); kennzeichnend für das ZiPP sind ferner
- die *freiwillige Befragung*, wobei die gezogene Stichprobe in Erwartung einer nicht unbeträchtlichen Nichtbeantwortung<sup>16</sup> bewusst sehr großzügig dimensioniert war
- Fragen zur *Unternehmensdemographie* (Organisationswechsel, Ein- und Austritt von Partnern, Änderung des *Versorgungsauftrags* etc.), die bei der KSE weitgehend fehlen
- es werden im Vergleich zur KSE mehr Details zu Einnahmen gefragt, insbesondere *auch* zu *Versorgungsleistungen* (z.B. Behandlungsfälle, Arbeitszeit der Ärzte und des nichtärztlichen Personals; Praxisöffnungszeit etc.), die bei der KSE ebenfalls weitgehend fehlen, was auch gilt für
- Fragen zur räumlichen und technischen *Ausstattung von Praxen* und
- aus ähnlichen Gründen (Aspekte der Versorgung und Vergütung) enthält das ZiPP auch anders als die KSE Fragen nach (z. Zt. noch nicht so relevanten aber evtl. wichtiger werdenden) neuen „selektivvertraglichen Instrumenten“ und zu erwähnen sind ferner
- Fragen zum *Vermögen* und dessen *Finanzierung* sowie zu getätigten und für erforderlich gehaltenen *Investitionen* und dass ein
- *Testat des Steuerberaters* bei den erhobenen finanziellen Daten verlangt wird.<sup>17</sup> Zu beachten ist schließlich auch, dass das ZiPP im Unterschied zur KSE, bei der
- *Regionalisierung* der Daten zu Versorgung, Kosten und Erträge weiter geht (mehr differenziert) als die KSE und auch anders gebildete regionale Bezugseinheiten verwendet<sup>18</sup>
- die Möglichkeit ZiPP-Angaben mit Abrechnungsdaten (VDX-Datei) zu verknüpfen was neben zusätzlichen Plausibilisierungsmöglichkeiten vor allem ein insgesamt erheblich größeres Analysepotential impliziert als es bei der KSE gegeben ist, dabei ist zu beachten, dass
- die beabsichtigte Verwendung nicht (oder zumindest nicht allein) in der Information der „Allgemeinheit“ mit „unbearbeiteten“ Daten, so wie sie erhoben worden sind besteht, sondern das Interesse an Einnahmen, Kosten, Reinerträgen usw. besteht, wie sie *wären*, wenn man eine „Normpraxis“ bzgl. Versicherungsart der Patienten und

<sup>16</sup> oder „echte Antwortausfälle“ in der später eingeführten Terminologie

<sup>17</sup> Die Befragung besteht somit aus zwei Teilen, einem Teil A, der von den Ärzten auszufüllen ist und Angaben über die Praxis (Ausstattung, Personal, Arbeitszeiten, Behandlungsfälle, fachliche Spezialisierung etc.) erfragt und einen Teil B, der sich an die Steuerberater richtet. Dass bei der Befragung notwendig ein Steuerberater involviert sein muss hat sich im Zuge der Erhebung als (unerwartet) problematisch herausgestellt.

<sup>18</sup> Hier zeigt sich, dass eine Analyse der Zwecke einer Statistik sehr wichtig ist, denn diese Besonderheiten des ZiPP ergeben sich daraus, dass das ZiPP auch Daten liefern sollte für Honorarverhandlungen und Anpassung „Orientierungswert“ und dem einheitlichen Bewertungsmaßstab (EBM). Dabei sind ausdrücklich (§ 87 Abs. 2f SGB V) Hinweise auf regionale Besonderheiten und die Sicherstellung der vertragsärztlichen Versorgung von Bedeutung.

Arbeitszeit hätte; d.h. anders als bei der KSE ist beim ZiPP an bestimmte „Standardisierungen“ als wesentliches Element der Analyse und Ergebnispräsentation.

## 2. Daten zu den Stichproben (Auswahlverfahren, Erhebung, Hochrechnung)

In diesem und in dem nächsten Kapitel 3 werden weitere Details zu den beiden Befragungen mitgeteilt, die in dem grobem Überblick des Kapitels 1 noch so eingehend behandelt werden konnten. Im Kapitel 4 werden dann auch Konzeptions- und Auswertungsprobleme des ZiPP dargestellt, wie z.B. der Panelaspekt oder die Standardisierung, auf die im Kapitel 1 (dort im Abschn. 1.3) nur kurz hingewiesen werden konnte.

### 2.1 Bereichsabgrenzung und Auswahlgesamtheit

Der in der KSE 2007 einbezogene Bereich aus dem Abschnitt N der NACE betrifft Unternehmen und Arbeitsstätten (Praxen) der Bereiche (Wirtschaftszweige)

85.12 Arztpraxen (ohne Zahnarztpraxen)

85.12.1 Arztpraxen für Allgemeinmedizin und Praxen von praktischen Ärztinnen und Ärzten

85.12.2 Facharztpraxen (ohne Arztpraxen für Allgemeinmedizin)

85.13.0 Zahnarztpraxen

85.14.1 Praxen von psychologischen Psychotherapeutinnen und -therapeuten<sup>19</sup>

85.20.1 Tierarztpraxen

Der Bereich 85.14.1 gehört auch im ZiPP zur Zielgesamtheit. Zahnärzte und Tierärzte gehören zwar zur Gesamtheit in der KSE, nicht aber zu der des ZiPP. Auch die benutzte Auswahlgesamtheit ist nicht vergleichbar. Das bei den Statistischen Ämtern der Länder und des Bundes geführte statistische Unternehmensregister<sup>20</sup> (URS) legt die Auswahlgesamtheit der KSE fest, d.h. aus ihm wird die Stichprobe der Erhebungseinheiten gezogen. Es dient – wie bereits erwähnt – u. a. als Steuerungsinstrument bei Erhebungen, Auswahl- und Hochrechnungsrahmen sowie als Quelle für sonst speziell zu erhebenden Strukturdaten. Es ist geplant immer mehr Primärerhebungen durch eine registerbasierte Statistik (also durch Sekundärdatennutzung) zu ersetzen.<sup>21</sup> Beim Aufbau des URS verwendete Datenquellen sind vor allem Angaben der Arbeitsverwaltung (sozialversicherungspflichtige Beschäftigte) und selber aus eigenen Primärerhebungen gewonnene Daten der amtlichen Statistik, seltener aber – speziell im Falle von Ärzten – Daten der Finanzverwaltung (Umsatzsteuer), die sonst für den Aufbau des URS von großer Bedeutung ist.<sup>22</sup> Zur Zeit der Stichprobenziehung (April 2008) befanden sich im URS etwa 159.000 Praxen (vgl. Tab. 1).<sup>23</sup>

Die entsprechende Auswahlgesamtheit beim ZiPP ist das Bundesarztregister (BAR) und dieses ist notwendigerweise vollständig, weil die Zulassung mit der Eintragung verbunden ist. Auch eine ruhende Zulassung bei einer Einheit wird im BAR vermerkt, und das würde bedeuten, dass diese Einheit auch nicht für die Stichprobe gezogen werden kann.

<sup>19</sup> Kinder- und Jugendlichenpsychotherapeuten sind nur berechtigt, Patienten zu behandeln, die das 21. Lebensjahr noch nicht vollendet haben, während Psychologische Psychotherapeuten alle Patienten behandeln dürfen, also auch Kinder und Jugendliche. Die Schreibweise „Praxen von psychologischen Psychotherapeuten“ umfasst die Praxen von Psychologischen Psychotherapeuten und Kinder- und Jugendlichenpsychotherapeuten.

<sup>20</sup> auch genannt: Unternehmensregister für statistische Verwendungszwecke (URS)

<sup>21</sup> Entsprechende Überlegungen sind auch im Zusammenhang mit Ärzteregeatern relevant. Im ZiPP ist vorgesehen, zumindest einen Teil der Erhebung aus dem Bundesarztregister abzuleiten. Zu dieser Thematik (insbes. aus Sicht des URS) vgl. auch R. Sturm, Konzept einer registerbasierten Wirtschaftsstatistik, in: Wirtschafts- und Sozialstatistisches Archiv, Bd. 4 (2010), S. 115 – 139.

<sup>22</sup> Da Ärzte nicht umsatzsteuerpflichtig sind entfallen hier entsprechende Informationen im URS. Das StBA konnte daher auch nur nach der Zahl der sozialversicherungspflichtigen Beschäftigten (als Indikator für die Praxisgröße) schichten und nicht nach dem Umsatz.

<sup>23</sup> Es wird unterschieden zwischen Texttabellen (Tab. 1, Tab. 2, ...) und z.T. wesentlich umfangreichere ergänzende Tabellen im Anhang (Tab. A1, Tab. A2, ...).

**Tabelle 1**

Daten zum Auswahlrahmen (Vergleich KSE/ZiPP)

	URS	BAR
Anzahl der Registereinheiten insgesamt (alle bei der KSE erhobenen Bereiche)	etwa 159.000 (vgl. auch Tab. 2)	
<b>nur Arztpraxen</b> (ohne Zahnarztpraxen) darunter reine Privatpraxen	Zahlen für diese Differenzierung nicht feststellbar	nicht enthalten
Praxen mit Zulassung zur kassenärztlichen Versorgung		94.594 (vgl. Tab. A1 im Anhang)

URS = Statistisches Unternehmensregister (KSE), BAR = Bundesarztregister (ZiPP)

Die Übersicht 3 fasst die wichtigsten erhebungstechnischen Unterschiede der beiden Erhebungen zusammen.

Beide Erhebungen, die KSE und das ZiPP sind geschichtete Stichproben und in beiden Erhebungen wurde aus Schichten, in denen nur wenige Praxen enthalten waren 100% der Einheiten erfasst. Für die Zuordnung einer Praxis zu einer der 108 Schichten (aufgrund von 36 Abrechnungsgruppen (Fachgruppen) und drei Regionstypen) ist beim ZiPP vor allem die Abrechnungsgruppe nicht immer unproblematisch. Die Zuordnung der Praxen erfolgt jeweils nach dem (ersten) Zulassungsfachgebiet des in der Praxis tätigen Arztes bzw. der dort tätigen Vertragsärzte. Damit soll eine *tätigkeitsorientierte Adressierung* sichergestellt werden, die garantieren soll, dass ein zugelassener Allgemeinarzt mit Weiterbildung Chirurgie als „Allgemeinarzt“ und nicht als „Chirurg“ angeschrieben wird, wenn er nicht als Chirurg zugelassen ist. Die Liste der 36 Fachgruppen, nach denen im ZiPP gegliedert wurde ist den Tab. A1 und A2 im Anhang zu entnehmen.<sup>24</sup>

Durch die Auswahl der Ärzte nach Zulassungsfachgebieten, die Beschränkung auf Praxen zugelassener Vertragsärzte, also die Nichtberücksichtigung ermächtigter und angestellter Ärzte, die Zusammenfassung von Praxen mit Ärzten unterschiedlicher Zulassungsfachgebiete in separaten Abrechnungsgruppen<sup>25</sup>, die Beschränkungen auf Praxen mit Zulassung vor dem 1.1.2008 und die zur Verfügung stehende Auswahlbasis (Bundesarztregister, Stand 31.12.2008) resultieren für einige Abrechnungsgebiete/-gruppen vergleichsweise kleine Praxiszahlen bzw. erhebliche Abweichungen zu den zum Teil nach Weiterbildungsordnung differenzierten Arztstatistiken. Zur Gliederung nach Arztgruppen und der entsprechenden Differenzierung der Fragebögen finden sich im Tabellen-Anhang einige weitere Angaben.

<sup>24</sup> Einige ergänzende Bemerkungen (die Nummern 1, 16 und 35 beziehen sich auf Tab. A2):

(1) „Hausärzte“ umfassen auch fachübergreifende Gemeinschaftspraxen von Allgemeinmedizinern, prakt. Ärzten und Hausarzt-Internisten, in Gruppe (16) Pneumologen sind auch Fachärzte für Lungenheilkunde enthalten und (35) enthält nicht fachübergreifende Gemeinschaftspraxen von Allgemeinmedizinern, prakt. Ärzten und Hausarzt-Internisten (die nach dem oben Gesagten in der Gruppe 1 gezählt werden).

<sup>25</sup> Das betrifft insbesondere die fachinternistischen Schwerpunktpraxen, welche in die Gruppe „sonstige Fachinternisten“ fallen, wenn sie mehrere Schwerpunkte aufweisen, dagegen in die Gruppe fachübergreifender Gemeinschaftspraxen, wenn sie Praxen mit Hausärzten oder Ärzten anderer Fachrichtungen betreiben.

## Übersicht 3

Abgrenzung des Erhebungsbereichs und Auswahlgesamtheit

	KSE (StBA)	ZiPP (ZI)
Erhebungs-gesamtheit	Erhebungsgesamtheit definiert auf Basis der EU-einheitlichen Systematik der Wirtschaftszweige NACE Rev.1.1, nur Praxen mit einem Mindestumsatz von 12.500 € im Erhebungsjahr <sup>a)</sup> ; ausgeschlossen auch Praxen von Existenzgründern (im Sinne von §7g Einkommensteuergesetz) <sup>b)</sup>	Definiert mit den Voraussetzungen für die Zulassung zur kassenärztlichen Versorgung gem. Zulassungsordnung für Vertragsärzte (Ärzte-ZV) <sup>c)</sup> , dies sind zugleich Voraussetzungen für Eintragung in das (Bundes-) Arztregister der KVen bzw. KBV; Praxen als Erhebungseinheiten
Stichprobe	Geschichtete Stichprobe <sup>d)</sup> (offenbar disproportionale Aufteilung) <sup>e)</sup>	Geschichtete Stichprobe (108 Schichten, vgl. Übers. 2)
Auswahl-gesamtheit und Auswahl-einheit	Unternehmensregister des StBA (mit ca. 90.000 Einheiten, darunter auch reine Privatpraxen), Auswahl-einheit (und zugleich Erhebungs- und Darstellungseinheit) ist die einzelne Praxis	Bundesarztregister (BAR) der KBV (Kassenpraxen mit Strukturdaten). Befragung von Praxen (als Auswahl-einheiten). Stichprobenplanung bei Erwartung von nur ca. 20% Rücklauf.
Durchführung der Befragung	Stichprobenziehung April 2008, Fragebogenversand 26.9.2008 (Befragung für Berichtsjahr 2007)	Stichprobenziehung aus dem BAR 32.700 Versand in 3 Wellen (März/ April 2010); Frist f. Rücklauf wurde verlängert
erzielter Rücklauf	6.237 von 7.843 befragten Einheiten (79,5%), vgl. Tab. 2	4.661 bearbeitbare Eingänge, davon 4.636 verwertbar
Antwort-ausfälle	vgl. Tab. 3. zur Systematik der Antwortausfälle und der Unterscheidung in echte und unechte Antwortausfälle vgl. Abb. 1, weitere Angaben unter Abschn.	
Hochrechnung	freie Hochrechnung (Hochrechnungsfaktor ist der Kehrwert des Auswahl-satzes der Schicht also $N_s/n_s$ , korrigiert um die "echten" Antwortausfälle)	freie Hochrechnung, auch um beträchtliche und zudem ungleichmäßig verteilte "echte" Antwortausfälle auszugleichen
Ergebnis	veröffentlicht am 5.8.2009 in Fachserie 2	veröffentlicht am 16.3.2012 als erster Jahresbericht

- a) Diese Abschneidegrenze galt schon bei der KSE 2003. Angaben für „Tätige Personen“ sind stichtagsbezogen.  
 b) Befreiung im Gründungsjahr und in den beiden folgenden Jahren wenn Umsatz unter 500.000 € liegt.  
 c) Dazu gehören u.a. Approbation, Weiterbildung (§3 Ärzte-ZV).  
 d) Schichtung in der ersten Stufe nach Gebietszugehörigkeit (Früheres Bundesgebiet und Neue Länder und Berlin-Ost), in der zweiten Stufe nach den Wirtschaftszweigen (Facharztbezeichnungen bei den Ärzten) und in der dritten Stufe nach der Zahl der sozialversicherungspflichtig Beschäftigten (insges. 254 Schichten).  
 e) D.h. ungleiche Auswahl-sätze  $n_s/N_s$  für die Schichten  $s = 1, 2, \dots, S$  (hier  $S = 254$ ), denn: „Die Auswahl-wahrscheinlichkeit ist für einnahmestarke Erhebungseinheiten im Allgemeinen größer als die für einnahmeschwächere Einheiten.“

Ein besonders wichtiger Unterschied ist, dass in der KSE auch reine Privatpraxen enthalten sind, nicht nur Praxen von zugelassenen Kassenärzten, wie im ZiPP und dass Existenzgründungen in der KSE grundsätzlich ausgeschlossen sind.

## 2.2 Stichprobenplanung und Durchführung der Stichprobenziehung

Bei der ZiPP-Stichprobe stand man vor der Frage, die Struktur der Grundgesamtheit nicht ausreichend genau zu kennen, um mehr als nur eine grobe Abschätzung des notwendigen Stichprobenumfangs  $n^*$  vornehmen zu können. Dabei wurde jede Stichprobe aus jeder Schicht (des Umfangs  $N_k$ ) als eine Stichprobe für sich betrachtet.<sup>26</sup> Es ist unmittelbar einsichtig, dass der erforderliche Stichprobenumfang  $n_k^*$  der Stichprobe aus der  $k$ -ten Schicht von drei Fak-

<sup>26</sup> Das bedeutet, dass man die entsprechenden Formeln (zur Abschätzung von  $n^*$ ) für die „einfache“ Stichprobe benutzt hatte um eine gewünschte Genauigkeit für die betreffende Schicht  $k$  ( $k = 1, \dots, K$ ) sicherstellen zu können. Diese Formeln unterscheiden sich von denen für eine geschichtete Stichprobe. Auf die Problematik der beiden Betrachtungsweisen wird oben im laufenden Text weiter eingegangen.

toren abhängt, von

1. der gewünschten *Genauigkeit*  $e$  (bzw. genauer „Ungenauigkeit“, gemessen als halbe Breite des symmetrischen zweiseitigen Konfidenzintervalls für den Grundgesamtheitsmittelwert  $\mu$ , das sich auf Basis des Stichprobenmittelwerts  $\bar{x}$  bestimmt, also mit  $e = z_\alpha \hat{\sigma}_{\bar{x}}$  (hierin ist die Größe  $\hat{\sigma}_{\bar{x}}$  der „Stichprobenfehler“);
2. der gewünschten *Sicherheit*  $1 - \alpha$  (z.B.  $1 - \alpha = 0,9$  bedeutet 90% „Sicherheit“, oder  $1 - \alpha = 0,95$  ist 95% usw.), was wiederum die Größe  $z_\alpha$  bestimmt, die umso größer ist je größer die geforderte Sicherheit ist, so ist z.B.  $z_\alpha = 1,645$  bei 90% und  $z_\alpha = 1,960 \approx 2$  bei 95%, womit im Folgenden gerechnet wird), und
3. von der *Homogenität der Grundgesamtheit*, die in der Varianz  $\sigma^2$  (bzw.  $\sigma_k^2$  der k-ten Schicht) zum Ausdruck kommt (es spielt dabei natürlich auch eine Rolle, an welcher Variable  $x$  man sich bei der ganzen Betrachtung orientieren will).<sup>27</sup>

Der Stichprobenfehler ist dann gegeben mit  $\hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}_k}{\sqrt{n_k}} \sqrt{\frac{N_k - n_k}{N_k - 1}} \approx \frac{\hat{\sigma}_k}{\sqrt{n_k}} \sqrt{1 - \frac{n_k}{N_k}}$  und die

Bestimmung von  $n_k^*$  läuft darauf hinaus, die Gleichung  $e = z_\alpha \hat{\sigma}_{\bar{x}}$  nach  $n$  aufzulösen. Das Problem dabei ist, dass dieses Vorgehen Kenntnis der Varianz  $\sigma^2$  der Variable  $x$  in der Grundgesamtheit voraussetzt, die bei Planung der ersten ZiPP-Befragung nicht gegeben war. Geht man von einem Merkmal aus, das nur zwei Ausprägungen (etwa ja/nein) hat und gilt es den Anteilswert (Prozentsatz der ja-Antworten)  $\pi_k$  in der Grundgesamtheit aufgrund des Anteils  $\hat{\pi}_k$  in der Stichprobe<sup>28</sup> zu schätzen, so vereinfachen sich die Formeln erheblich, weil die jetzt an die Stelle der Varianz  $\sigma_k^2$  tretende Größe  $\pi_k(1-\pi_k)$  einen maximalen Wert von  $\frac{1}{4}$  nicht übersteigen kann. Um „auf der sicheren Seite“ zu sein, kann man deshalb einheitlich (für alle Schichten)  $\frac{1}{4}$  vorgeben. Rechnet man zudem einheitlich mit  $z_\alpha = 2$  (also einer Sicherheit von gut 95%) und einem Fehler von  $e = 0,08$  so erhält man die vom ZI angewendete Formel

$$(1) \quad n_k^* \geq \frac{1}{e^2 + \frac{1}{N_k}} = \frac{N_k}{0,0064 \cdot N_k + 1} \cdot 29$$

Die Planung ging davon aus, dass eine Gesamtstichprobe von weit über 7.000 Praxen angestrebt werden sollte. Die erwartete unvollständige Beteiligung verlangte einen entsprechend größer dimensionierten Fragebogenversand (in vielen Fällen wurde mit einem fünffachen Umfang gerechnet).<sup>30</sup>

Das Problem der Vorgehensweise gem. Gl. (1) bei der Stichprobenplanung ist, dass der Auswahlsatz  $n_k/N_k$  von  $N_k$  abhängig ist (großes  $N_k$ , wie z.B. bei Hausärzten bedeutet kleiner Auswahlsatz von nur 1,3%, kleines  $N_k$ , wie etwa bei Endokrinologen dagegen ein großer Auswahlsatz von 81,5%)<sup>31</sup> und es hängt von der Anzahl  $K$  der Schichten ab, nach denen differenziert wird wie groß dann im Endeffekt die Gesamtstichprobe  $n = \sum n_k$  ist. Es ist auch leicht zu sehen, dass mit zunehmenden  $N_k$  die Größe  $n_k^*$  zu einem festen Wert von  $1/e^2 = 1/0,0064 = 156$  strebt (und entsprechend der Auswahlsatz gegen Null).

Die Formeln für die geschichtete Stichprobe sind hier jedoch kein Ausweg, weil sie die Abschätzung von  $n$  für die Gesamtstichprobe voraussetzen (die natürlich ein  $n$  von 156 ergäbe für alle Schichten zusammen genommen!) und nur die Aufteilung von  $n$  in die Teilstichproben  $n_1, n_2, \dots, n_K$  (so dass  $n = n_1 + n_2 + \dots + n_K$  ist) betreffen. Außerdem geht es bei ihnen um die

27 Es ist unmittelbar einleuchtend, dass ein Stichprobenumfang von  $n = 1$  ausreichen würde, um  $\mu$  aus der Stichprobe mit Sicherheit ( $1 - \alpha = 1$ ) genau zu bestimmen, wenn alle Einheiten der Grundgesamtheit bezüglich  $x$  gleich wären (also  $\sigma^2 = 0$  wäre). Je inhomogener die Grundgesamtheit ist (je größer  $\sigma^2$  ist) desto größer muss  $n$  sein.

28 jeweils bezogen auf die k-te Schicht

29 Außerdem sollte auch für kein  $k$   $n_k$  kleiner als 30 sein.

30 Wie Tab. A1 im Anhang zeigt, ist es im Schnitt eine etwa 3,6 fache (358,76%) Zahl von versandten Fragebögen (angeschriebene Praxen AP) im Verhältnis zur kalkulierten Stichprobe (KS).

31 Die Prozentsätze errechnen sich aufgrund der Angaben von Tab. A 1 im Anhang.



Genauigkeit der Schätzung des Gesamtmittwerts  $\mu = \sum \frac{N_k}{N} \mu_k$  und nicht der Schichtmittelwerte  $\mu_1, \dots, \mu_k$ , die im Rahmen des

ZiPP natürlich von besonderem Interesse sind.

Mit Kenntnis der Varianzen  $\sigma_k^2$  der Variable(n)  $x$  (bzw.  $x_1, x_2, \dots$ ) nach einer ersten ZiPP-Stichprobe ließe sich natürlich eine neuerliche Stichprobe sehr viel besser planen, als nur mit Gl. (1). Außerdem sind diese Größen, also  $n_1, n_2, \dots, n_k$  – wie obige Formeln zeigen – wichtig zur Abschätzung des (absoluten) Stichprobenfehlers  $\hat{\sigma}_{\bar{x}}$ , bzw. des relativen Stichprobenfehlers  $\hat{\sigma}_{\bar{x}}/\bar{x}$  (vgl. Abschn. 2.5).

Die Stichprobenplanung des Bundesamts erfolgte im Falle der KSE – wie auch sonst in der amtlichen Statistik – offenbar nach dem Prinzip der „vergleichbaren Präzision“ (d.h. vergleichbar großer Stichprobenfehler), das verlangt, die Auswahlsätze so zu planen, dass eine weitgehend gleiche vorgegebene Genauigkeit für alle Teilgesamtheiten (z.B. Bundesländern) garantiert ist. Die Einzelheiten des offenbar nicht unkomplizierten Vorgehens sind uns nicht bekannt.<sup>32</sup>

Die Durchführung der ZiPP-Stichprobe erfolgte in der Weise, das in der Zeit von 22. März bis 15. April 2010 31.833 niedergelassene Ärzte und Psychotherapeuten in drei Versandwellen<sup>33</sup> vom Deutschen Ärzte-Verlag angeschrieben wurden. Die Frist für eine Einreichung von ausgefüllten (einschl. Testat des Steuerberaters) Erhebungsunterlagen wurde bis zum 31. März 2011 verlängert. Gleichwohl gab es Antwortausfälle, mit denen allerdings angesichts der Kompliziertheit der Befragung zu rechnen war.

### 2.3 Fehlersystematik, Antwortausfälle (nicht-stichprobenbedingte Fehler)

Für die Beurteilung der Qualität einer Stichprobe wird gerne der *Auswahlsatz*  $n/N$  und die *Antwortquote* (oder *Rücklaufquote*)<sup>34</sup>  $m/n$  herangezogen.<sup>35</sup> Nach Angaben des StBA war der Auswahlsatz bei der KSE 2007 rund 5% und von diesen 5% gab es nur 20,5% Nichtbeantwortung (nonresponse) oder nicht verwertbare Beantwortung, so dass man von einer Rücklaufquote (in Prozent des Fragebogenversands) von 79,5% sprechen kann. Die entsprechenden Relationen sind beim ZiPP 9,38% gemessen an dem angestrebten Stichprobenumfang, bzw. 33,6% (= 31.833/94.594) gemessen am Fragebogenversand (vgl. Spalten 7 und 8 in Tab. 2 und die Zahlenangaben der Tab. A1 im Anhang), was den Auswahlsatz betrifft und 14,6% (= 4.661/31.833) beim Rücklauf im Vergleich zum Fragebogenversand.<sup>36</sup>

Ob der Auswahlsatz als Kriterium der „Güte“ oder „Repräsentativität“ gelten kann, ist mit Recht umstritten. Wie gut ein Auswahlsatz einer bestimmten Höhe ist, hängt von den bei Bestimmung des Stichprobenumfangs genannten Faktoren ab. Bei einer entsprechend homogenen

32 Offenbar führt das Auswahlverfahren zu ungleichen Auswahlätzen  $n_k/N_k$  die umgekehrt proportional zur Schichtstärke sind, denn das Bundesamt schreibt: „In Schichten, die quantitativ sehr stark besetzt waren, wurde unter Berücksichtigung des vorgegebenen Auswahlsatzes eine repräsentative Stichprobe von Praxen gezogen. Das führt dazu, dass in Erhebungsbereichen, die nur eine geringe Anzahl von Praxen in der Auswahlgesamtheit hatten, eine relativ hohe Anzahl von Praxen gezogen und befragt wurden, um ein repräsentatives Hochrechnungsergebnis zu erhalten. Entsprechend niedriger ist der Auswahlatz in Erhebungsbereichen mit einer hohen Anzahl von Praxen in der Auswahlgesamtheit.“

33 Vor allem um das Risiko einer Budgetüberschreitung zu minimieren, wenn sich die Rücklauferwartung als zu pessimistisch erweisen sollte.

34 Üblich ist auch der Begriff „Ausschöpfungsquote“.

35 Im Qualitätsbericht heißt es (zu den Zahlen vgl. Tab. 2) dazu: Im Ergebnis der Stichprobenziehung wurden bundesweit 7.843 Praxen befragt. Das entspricht einem Auswahlatz von 5%. Von den 7.843 befragten Praxen erhielt das Statistische Bundesamt 6.237 Fragebogen mit verwertbaren Angaben zurück, was einem Anteil von 79,5% entspricht (zu den Zahlen vgl. Tab. 2 Spalten 3 und 5).

36 Diese Quote (14,6%) wird auch „Realisierungsquote“ genannt (valide Antworten in Relation zum Fragebogenversand, bereinigt um unechte Antwortausfälle). Vgl. hierzu, Das PHF: eine Erhebung zu Vermögen und Finanzen privater Haushalte in Deutschland, Monatsberichte der Deutschen Bundesbank, Januar 2012, S. 29ff. Bei dieser Erhebung, die in vieler Hinsicht Ähnlichkeiten mit dem ZiPP hat (freiwillige Erhebung von finanziellen Angaben, ein Panel ist intendiert), war die Realisierungsquote 18,6%, was im Vergleich zu nationalen und internationalen Erfahrungen als niedrig eingestuft wurde.

Grundgesamtheit kann ein Auswahlatz von 5% oder 9% sehr groß sein, er kann aber bei einer inhomogenen Grundgesamtheit oder bei sehr großen Genauigkeitsanforderungen sehr klein sein. Es ist insbesondere nicht sinnvoll, „gefühlsmäßig“ von einem (geringen) absolutem Stichprobenumfang („nur ... Befragte“) auf „Nichtrepräsentativität“ zu schließen. Das gilt schon deshalb, weil unklar ist, was mit „Repräsentativität“, (was bezeichnenderweise kein Fachausdruck aus der Statistik ist) überhaupt gemeint ist.<sup>37</sup> Die Rolle des ziemlich brauchbaren Begriffs „Repräsentativität“ im nicht-statistischen Sprachgebrauch übernimmt der exakte und quantifizierbare Begriff „Stichprobenfehler“ (oben bereits erwähnt mit  $\hat{\sigma}_{\bar{x}}$  bzw.  $\hat{\sigma}_{\hat{\pi}}$ , also der geschätzten Standardabweichung der Stichprobenverteilung des arithmetischen Mittels  $\bar{x}$ , bzw. des Anteilswerts  $\hat{\pi}$ ) in der Fachterminologie.

**Tabelle 2**

Anzahl der befragten Praxen und Rücklauf der Fragebogen (Vergleich KSE/ZIPP)

KSE (StBA)					ZIPP (ZI)			
Wirtschaftszweig	Praxen in Auswahlgesamtheit	Stichprobenumfang*	(3) in % von (2)	Verwertbare Fragebögen	(5) in % von (2)	Praxen in Auswahlgesamtheit	versandte Fragebögen	Verwertbare Fragebögen
1	2	3	4	5	6	7	8	11
85.12	89.133	5.299	5,95	4.090	4,59	78.916	29.618	4.130
85.12.1	37.084	953 <sup>4)</sup>	2,57	709	1,91	39.916	4.336	598
85.12.2	52.049	4.346 <sup>5)</sup>	8,35	3.381	6,50	39.000	25.282	3.532
85.14.1	13.400	670	5,00	600	4,48	15.678	2.215	531
<b>insges.</b>	<b>102.533</b>	<b>5.969<sup>6)</sup></b>	<b>5,82</b>	<b>8.780</b>	<b>4,57</b>	<b>94.594<sup>3)</sup></b>	<b>31.833<sup>4)</sup></b>	<b>4.661</b>

nachrichtlich

1	2	3	4	5	6
85.13.0	46.922	1.458	3,11	1.230	2,62
85.20.1	8.718	416	4,77	317	3,64
<b>alle</b>	<b>158.173</b>	<b>7.843</b>	<b>4,96</b>	<b>6.237</b>	<b>3,94</b>

Für 85.13.0 (Zahnarztpraxen) und 85.20.1 (Tierarztpraxen) keine Entsprechung im ZIPP

\*Stichprobenumfang = Anzahl der versandten Fragebögen

Zum NACE-Code

85.12 Arztpraxen (ohne Zahnarztpraxen)

85.12.1 Arztpraxen für Allgemeinmedizin und Praxen von praktischen Ärztinnen und Ärzten

85.12.2 Facharztpraxen (ohne Arztpraxen für Allgemeinmedizin)

85.14.1 Praxen von psychologischen Psychotherapeutinnen und -therapeuten

Der Stichprobenfehler (Auswahlfehler) ist ein durch die Stichprobenziehung bedingter Fehler. Er besteht darin, dass bei einer Stichprobe nicht alle Einheiten in die Stichprobe einbezogen sind und er würde deshalb bei einer Vollerhebung nicht auftreten. Um ihn geht es, wenn es gilt, die „Repräsentativität“ zu beurteilen. Die Besonderheit einer Stichprobe (= Zufallsauswahl) ist, dass der Auswahlfehler ein Zufallsfehler ist<sup>38</sup>, also mit den Gesetzen der Wahrscheinlichkeitsrechnung zu quantifizieren ist.<sup>39</sup> Wir kommen hierauf in Abschn. 2.5 zurück.

37 Das „Alltagsverständnis“ dieses Begriffs ist meist, dass in der Stichprobe eine gleiche oder ähnliche Struktur (Verteilung bestimmter Merkmale) gegeben ist, wie in der Grundgesamtheit. Es ist bei genauerem Hinsehen ein sehr fragwürdiges Konzept, obgleich der „Hochrechnung“ ähnliche Vorstellungen zugrundeliegen. Ich habe in verschiedenen Schriften versucht, dies darzulegen. Vgl. Peter von der Lippe und Andreas Kladroba, Repräsentativität von Stichproben, in Marketing ZFP 24 (2002), S. 139 – 145 und P. v. d. Lippe, Wie groß muss meine Stichprobe sein, damit sie repräsentativ ist? DISKUSSIONSBEITRAG aus der Fakultät für WIRTSCHAFTSWISSENSCHAFTEN der UNIVERSITÄT DUISBURG – ESSEN, Campus Essen, Nr. 187, Februar 2011, als Download unter von-der-lippe.org/publikationen4.php.

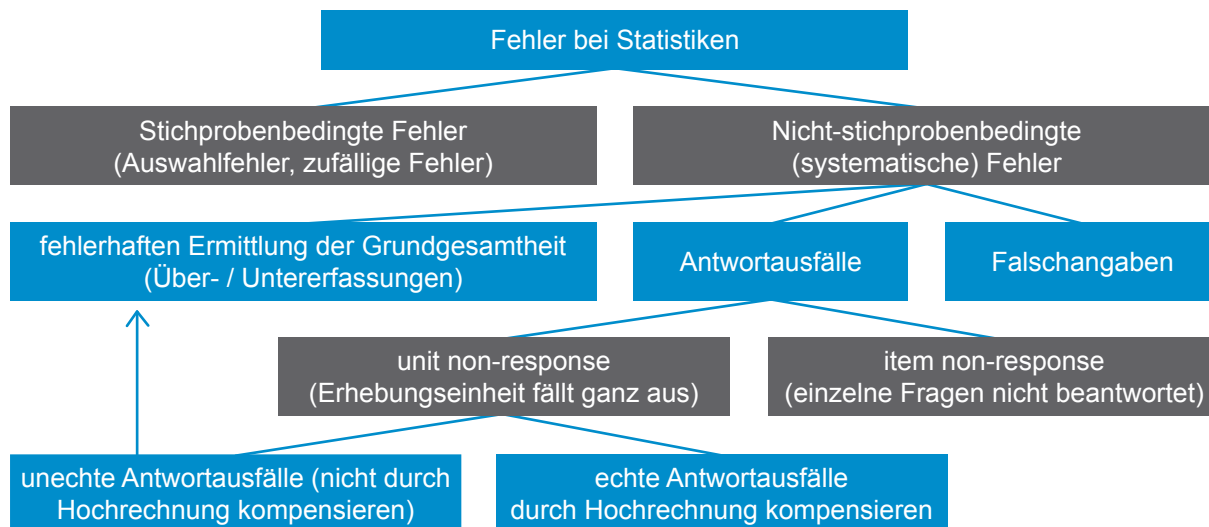
38 im Unterschied zu den meist nicht zufälligen „systematischen“ nicht-stichprobenbedingten Fehlern

39 Bei anderen Arten der Teilerhebung gilt das nicht. Der Vorteil, den Auswahlfehler quantifizieren zu können besteht nur wenn eine Zufallsauswahl vorliegt, also bei einer (echten) Stichprobe.

Anders verhält es sich mit der Nichtbeantwortung, gemessen an der Nonresponse-Quote<sup>40</sup>  $m/n$ , weil dies ein typischer nicht-stichprobenbedingter und in der Regel auch systematischer (nicht zufälliger, sich in der großen Zahl nicht ausgleichender) Fehler ist, der auch bei einer Vollerhebung auftreten kann (vgl. Abb. 1 zur Fehlersystematik).

### Abbildung 1:

Fehlersystematik bei statistischen Erhebungen



Wenn Nichtbeantwortung beträchtlich ist, kann argumentiert werden, dass nicht wirklich eine Stichprobe aus den  $N$  Einheiten der Grundgesamtheit, sondern faktisch nur aus den  $N_A < N$  auskunftsbereiten Einheiten der Grundgesamtheit gezogen wurde. Wegen der Kompliziertheit einer Panelbefragung ist nicht auszuschließen, dass in diesem Fall mehr mit Problemen der Nichtbeantwortung als bei einer Einmalerhebung zu rechnen ist (vgl. Abschn. 4.1).

Üblicherweise wird versucht, nicht-stichprobenbedingte Fehler durch Rückfragen bei den Befragten zu reduzieren und z.B. bei Nichtbeantwortung (echte Antwortausfälle) Erinnerungsschreiben zu versenden oder mögliche Falschangaben durch Einsatz von Plausibilitätskontrollen weitgehend zu erkennen und zu korrigieren. Das ist in beiden hier zu vergleichenden Erhebungen, der amtlichen KSE und dem ZiPP geschehen. Im Falle des ZiPP ergaben sich durch den Zugang zur VDX-Datei bessere Möglichkeiten der Plausibilitätskontrolle als sie im Falle der KSE gegeben waren.

Es ist wichtig, zu unterscheiden zwischen diesen echten Antwortausfällen, denen man häufig auch durch Hochrechnung zu begegnen versucht und den unechten Antwortausfällen. Letztere beruhen darauf, dass die Auswahlgesamtheit (d.h. das Unternehmensregister URS, bzw. das Bundesarztregister BAR) fehlerhaft (z.B. auch durch Fehlzugeordnungen von Praxen zu Arztgruppen und nicht berücksichtigten Umwidmungen) oder nicht up to date ist.

Untererfassungen (eine Einheit gehört zur Zielgesamtheit ist aber nicht im Auswahlrahmen), wie z.B. nicht erfasste Neugründungen sind kaum zu quantifizieren, wenn die tatsächliche Grundgesamtheit nicht aus anderen Gründen bekannt ist. Bei einer Übererfassung (Weiterführen von erloschenen Einheiten in der Auswahlgesamtheit)<sup>41</sup> würde ein Antwortausfall einer solchen Einheit „unecht“ sein. Er ist in der Hochrechnung nicht zu berücksichtigen, weil die betreffende Einheit ja eigentlich nicht zur Zielgesamtheit der Erhebung gehört.<sup>42</sup> Auf die Hoch-

<sup>40</sup> Leider oft Nonresponse-„Rate“ genannt, wie ja überhaupt heutzutage Worte wie „Rate“ oder „Quote“ ganz nach Belieben und ohne erkennbares Prinzip benutzt werden. Die „Nichtantwortquote“  $q$  (etwa  $q = 0,2$  also 20%) ist das Gegenstück zur „Aus schöpfungsquote“  $1 - q$  (das Verhältnis von auswertbaren Fragebögen zu den insgesamt versendeten Fragebögen).

<sup>41</sup> Eine Übererfassung der Grundgesamtheit liegt z.B. bei sog. „Karteileichen“ (ein von Statistikern „erfundener“ Begriff) vor.

<sup>42</sup> Daraus resultiert, dass bei Vorliegen unechter Antwortausfälle die hochgerechneten Ergebnisse der Erhebung in der Fallzahl (Anzahl der Einheiten) immer niedriger als die der Zielgesamtheit sind. Das Auftreten von unechten Antwortausfällen in den

rechnung wird im Abschn. 2.4 weiter eingegangen.

Es ist anzunehmen, dass unechte Antwortausfälle im ZiPP seltener als in der KSE sein dürften, weil das BAR wegen der damit verbundenen Zulassungsregelung notwendig vollständig und jeweils aktualisiert sein dürfte, dass aber umgekehrt mit echten Antwortausfällen bei einer freiwilligen Befragung, wie dem ZiPP mehr zu rechnen ist als bei einer amtlichen (und daher mit Auskunftspflicht versehenen) Erhebung, wie der KSE.<sup>43</sup>

### Tabelle 3

Geschätzte Anzahl der Antwortausfälle in der KSE  
(Angaben des Qualitätsberichts des Statistischen Bundesamts)

	unechte Antwortausfälle		echte Antwortausfälle		Antwortausfälle insgesamt	
	absolut	in %	absolut	in %	absolut	in %
Arztpraxen (ohne Zahnarztpraxen)	949	17,9	260	4,9	1.209 <sup>a)</sup>	22,8
Praxen von psycholog. PsychotherapeutInnen	56	8,4	14	2,1	70	10,5
insgesamt in KSE (auch Zahn- und Tierärzte)	1.260 <sup>b)</sup>	16,1	346	4,4	1.606	20,4

a) = 5.299 – 4.090 in Tab. 2

b) bei insgesamt 7.843 befragten Einheiten sind 1.260 genau 16,1%; ferner gilt  $1.606 = 7.843 - 6.237 = 1.260 + 346$

Zur Zahl der echten Antwortausfälle in Gestalt von „unit nonresponse“<sup>44</sup> beim ZiPP vgl. Tab. A1 im Anhang.<sup>45</sup> Echte Antwortausfälle können – *wenn* sie zufällig auftreten (verteilt sind) – im Rahmen der Hochrechnung „korrigiert“ werden, d.h. durch Erhöhung der Zahl der Erhebungseinheiten der gleichen Ziehungsschicht „(hin)eingeschätzt“ werden (Zuschätzung echter Antwortausfälle).<sup>46</sup> Üblicherweise stellen aber echte Antwortausfälle systematische Fehler dar, die nicht durch Hochrechnung ausgeglichen werden können, weil sie in verschiedenen Schichten auch unterschiedlich häufig sind, bzw. auch innerhalb einer Schicht bei einer bestimmten Praxisform häufiger auftreten können als bei einer anderen Praxisform.

## 2.4 Hochrechnung

Wenn eine Stichprobe als eine Miniaturausgabe der Grundgesamtheit aufgefasst werden kann, ist es der Zweck der sog. „Hochrechnung“, eine Anzahl  $n_i$  von Einheiten (Praxen), für die die Merkmalsausprägung  $x = x_i$  gilt, wieder auf die Dimension der Grundgesamtheit zurück zu vergrößern. Ist der Auswahlsatz 5%, also  $1/20$  so liegt es nahe, zu sagen  $n_i = 300$  Einheiten in der Stichprobe entsprechen  $N_i = 300 \cdot 20 = 6.000$  Einheiten der Grundgesamtheit, also  $n_i$  im Sinne der „freien Hochrechnung“ mit dem reziproken Auswahlsatz zu multiplizieren. Man kann die Hochrechnung auch nutzen um bestimmte Häufigkeiten als „Eckwerte“ an die Struktur der Grundgesamtheit anzupassen. So wurde z.B. gefunden, dass 80,9% der am ZI-Praxis-Panel (ZiPP) 2010 teilnehmenden Praxen Einzelpraxen waren, d.h. es war  $n_E/n = 0,809$ , während es im Bundesarztregister (BAR) nur 78,7% (oder  $N_E/N = 0,787$ ) der Arztpraxen sind, so dass Einzelpraxen „überrepräsentiert“ (und damit Gemeinschaftspraxen unterrepräsentiert<sup>47</sup>) zu sein

Schichten bewirkt eine Verstärkung der Merkmalsstreuung und damit ein Anwachsen von durch die zufällige Auswahl der Stichprobeneinheiten bewirkten Schätzfehlern.

43 Bei einer amtlichen Erhebung liegen echte Antwortausfälle nur dann vor wenn Praxen ihrer Auskunftspflicht nicht oder nicht ordnungsgemäß nachkommen.

44 Neben Antwortausfällen auf der Ebene der Einheiten (unit non-response) gibt es auch Ausfälle auf der Ebene einzelner Fragen (Merkmale), die jedoch weniger bedeutsam sind wenn sie nicht zur Unbrauchbarkeit der Angaben insgesamt führen.

45 Man erhält die Zahl der „echten“ Antwortausfälle, wenn keine unechten Antwortausfälle vorliegen als Differenz zwischen Fragebogenversand und Fragebogenrücklauf.

46 Diese Vorgehensweise ist immer dann verzerrungsfrei, wenn das Auftreten der echten Antwortausfälle innerhalb der Schicht als Zufallsereignis angesehen werden darf.

47 Die entsprechenden Anteile für Gemeinschaftspraxen sind 19,1% in der Stichprobe und 21,3% im BAR.

scheinen.<sup>48</sup> Dies kann durch einen Faktor („Gewicht“, oder auch „Designgewicht“) von 0,9728 „korrigiert“ werden, so dass man mit  $80,9 \cdot 0,9728 = 78,7\%$  den hochgerechneten Anteil der Einzelpraxen erhält.<sup>49</sup> Das „Gewicht“ 0,9728 ergibt sich aus

$$0,9728 = \frac{0,787}{0,809} = \frac{N_E/N}{n_E/n}$$

Die Hochrechnung führte im ZiPP vor allem zu einer „Höhergewichtung“ der Anzahl der Praxen von Allgemeinmedizinern („Hausärzten“), deren gewogene Anzahl sehr viel größer ist als die Anzahl der an der Stichprobe teilnehmenden Praxen (also die „ungewogene“ Anzahl vor der Hochrechnung). Entsprechend verringert wurde das Gewicht der anderen Praxen. Tab. 4 gibt hierfür einige Beispiele.

Wenn Gemeinschaftspraxen deshalb „unterrepräsentiert“ sind, weil hier vermehrt echte Antwortausfälle auftraten (also die geplante Stichprobe zwar einen korrekten Anteil von 21,3% wie im BAR vorsah, sich von den angeschriebenen Praxen aber nur 99,7% beteiligten, so dass man einen Anteil von nur 19,1% statt 21,3% in der Stichprobe erhielt) kann durch diese „Hochrechnung“ auch der Antwortausfall ausgeglichen werden.

#### Tabelle 4

Hochgerechnete und noch nicht hochgerechnete Häufigkeiten im ZiPP  
(Angaben zur Anzahl der Praxen 2008 im ZiPP Jahresbericht 2010 des ZI [Tab. 26])  
entsprechende Angaben für die amtliche KSE sind nicht verfügbar

Arztgruppe	vor <sup>a)</sup>	nach <sup>b)</sup>	Arztgruppe	vor <sup>a)</sup>	nach <sup>b)</sup>
Alle Praxen	4.172	4.172	Hals-Nasen-Ohren-Heilkunde	249	119
Allgemeinmedizin <sup>c)</sup>	319	1.663	Orthopädie	228	138
Chirurgie	179	94	Psychotherapie	447	704
Gynäkologie	415	314	Urologie	164	77

- a) vor der Hochrechnung (ungewichtet)  
b) nach der Hochrechnung (gewichtet)  
c) und Innere Medizin (hausärztlich)

Sowohl was Hochrechnung im Sinne einer

- Minimierung der Differenz zwischen der Struktur (bzw. Häufigkeitsverteilung) bestimmter Merkmale in der Stichprobe und der entsprechenden Struktur in der Grundgesamtheit betrifft, als auch was
- die Berücksichtigung (Korrektur) von Nichtbeantwortung durch Hochrechnung betrifft,

ist zu bedenken, dass mit der freien Hochrechnung eine relativ einfache Methode angewendet wurde, was einige weitere Anmerkungen zur Nichtbeantwortung in Abschnitt 4.1 geboten erscheinen lässt.<sup>50</sup>

Man findet in Veröffentlichungen über Statistiken selten Zahlenangaben, wie sie sich direkt aus der entsprechenden Erhebung, vor Durchführung einer Hochrechnung ergaben. Weil eine Hochrechnung im Prinzip darauf hinausläuft, Angaben von einigen Befragten stärker und von anderen Befragten weniger zu „gewichten“ sind nicht nur Häufigkeiten berührt, sondern es bleiben natürlich auch Mittel- und Anteilswerte, sowie Varianzen und andere beschreibende

48 Zahlenangaben aus dem Jahresbericht 2010 zum ZI-Praxis-Panel, S. 18f.

49 und mit  $100 - 80,9 \cdot 0,9728 = 21,3\%$  den hochgerechneten Anteil der Gemeinschaftspraxen

50 Sowohl hinsichtlich der Hochrechnung als auch der Methoden im Umgang mit Nichtbeantwortung (echte Antwortausfälle) ist die bereits zitierte Arbeit, Das PHF: eine Erhebung zu Vermögen und Finanzen privater Haushalte in Deutschland, Monatsberichte der Deutschen Bundesbank, Januar 2012, S. 29ff. von sehr großem Wert. Bei dieser Erhebung wurden die Randverteilungen an die Struktur des Mikrozensus durch eine entsprechende Gewichtung angepasst, und es werden (was sehr selten ist) Häufigkeiten vor und nach Hochrechnung (Gewichtung) mitgeteilt. Man findet in dem Aufsatz außerdem eine sehr kurze und klare Beschreibung der Methoden der „Imputation“ von Antworten im Fall von nonresponse.

Kennzahlen (Statistiken) davon nicht unberührt.

## 2.5 Stichprobenfehler, Genauigkeit und „Repräsentativität“

Der relative Stichprobenfehler (oder Standardfehler)  $\hat{\sigma}_{\bar{x}}/\bar{x}$  gilt als Gütekriterium (Maß der Genauigkeit, bzw. der stichprobenbedingten Fehler) und ein möglichst geringer Wert für  $\hat{\sigma}_{\bar{x}}$  ist eine Zielgröße bei der Stichprobenplanung (wie bereits erwähnt hat sich die Planung der ZiPP-Stichprobe nicht an den Fehler [oder besser: an eine gewünschte Genauigkeit] des Gesamtmittelwerts  $\bar{x}$ , sondern an den Schichtmittelwerten  $\bar{x}_k$ ,  $k = 1, \dots, K$  der  $K = 108$  Schichten orientiert). Aus der bereits genannten Formel

$$(2) \quad \hat{\sigma}_{\bar{x}} = \frac{\hat{\sigma}_k}{\sqrt{n_k}} \sqrt{\frac{N_k - n_k}{N_k - 1}}$$

geht hervor, dass der Stichprobenfehler bei gegebenen Schichtumfängen  $N_k$  und Varianzen  $\sigma_k^2$  von den Umfängen der Teilstichproben, und den Umfängen  $n_k$  der Stichproben abhängt. Ein relativ großer Wert für  $\hat{\sigma}_{\bar{x}_k}^2$  kann daher rühren, dass  $\hat{\sigma}_k^2$  groß ist, die Schichtbildung also z.T. nicht zu hinreichend homogenen Schichten geführt hat,<sup>51</sup> oder dass der Stichprobenumfang  $n_k$  für die betreffende Schicht  $k$  relativ klein war.<sup>52</sup>

Nach Stichprobenziehung lässt sich  $\hat{\sigma}_k^2$  (als Schätzung von  $\sigma_k^2$ ) bestimmen und damit auch  $\hat{\sigma}_{\bar{x}_k}^2$  für die Schichten  $k = 1, \dots, K$ . Der (geschätzte) Stichprobenfehler für ein Gesamttaggregat  $\bar{x}$ , also  $\hat{\sigma}_{\bar{x}}^2$  kann größer oder kleiner sein als der Stichprobenfehler für die Teilaggregate  $\bar{x}_k$ , also die  $K$  Größen  $\hat{\sigma}_{\bar{x}_k}^2$ .<sup>53</sup> In Tab. 5 sind einige relative Stichprobenfehler wiedergegeben, wie sie vom StBA für die KSE mitgeteilt wurden und in Tab. 6 entsprechende Angaben zu den Praxisaufwendungen 2008 gem. ZiPP für alle Ärzte und für die in der Stichprobe besonders stark vertretenen Fachgruppen.

### Tabelle 5

Relativer Stichprobenfehler in Prozent bei Kernindikatoren der Kostenstrukturerhebung  
Teilgesamtheit Arztpraxen (Bereich Ärzte = 85.12)  
(Angaben des Qualitätsberichts des Statistischen Bundesamts)

Indikator	85.12 (mit MVZ)*
x <sub>1</sub> Gesamteinnahmen je Praxisinhaber	0,86
x <sub>2</sub> Gesamtaufwendungen je Praxisinhaber	0,94
x <sub>3</sub> Reinertrag je Praxisinhaber	1,08

\* Für die relativen Stichprobenfehler (in %) bei 85.12 ohne (statt oben mit) MVZ gibt das StBA genau die gleichen Werte an. Die relativen Standardfehler sind nach diesen Angaben jedoch bei Praxen von psychologischen Psychotherapeuten mit 2 - 2,4% (und noch mehr bei Tierarztpraxen mit 4,21 bis 4,62%) größer.

51 Ein Beispiel wäre der große Stichprobenfehler bei „Sonstigen Praxen“ (s) mit 15,2% (bei  $n_s = 109$ ).

52 Ein Beispiel hierfür wäre der große Stichprobenfehler bei Nephrologen im ZiPP in Höhe von 27,2% bei einem Stichprobenumfang von nur  $n_n = 9$ .

53 Er ist insbesondere kein Mittelwert der Stichprobenfehler der Teilaggregate. Die Ergebnisse des ZiPP (bei weiterer Betrachtung von Zahlen nach Art von Tab. 5) zeigen deutlich, dass der Stichprobenfehler für das Gesamttaggregat kleiner als die meisten Stichprobenfehler für die Teilaggregate sein kann.

**Tabelle 6**

Relativer Stichprobenfehler in Prozent bei Praxisaufwendungen 2008 (ZiPP)  
(Angaben des Jahresberichts 2010 des ZI für das ZI-Praxis-Panel, Tab. 29 S. 57)

Gruppe	$n_k$ bzw. $n$	in vH*
Alle Praxen	4.172	2,4
Allgemeinmedizin	319	3,8
Gynäkologie	415	2,8
Psychotherapie	447	3,2

\* Die Angaben des ZI betreffen den Stichprobenfehler (Standardfehler ohne Endlichkeitskorrektur).

In der zitierten Quelle von Tab. 6 sind auch die Bestandteile des Standardfehlers  $\hat{\sigma}^2$ ,  $n$  und  $\bar{x}$  zu entnehmen. Für „Alle Praxen“ ist dabei angegeben (geschätzte) Standardabweichung  $\hat{\sigma}^2 = 29.8011$ ,  $n = 4.172$  und  $\bar{x} = 188.426$ . Daraus errechnet sich

$$\hat{\sigma}_{\bar{x}}^{(o)} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{298.011}{\sqrt{4.172}} = 4,614$$

und dividiert man dies durch  $\bar{x}$  so erhält man 0,024486 also die angegebenen 2,4%. Man kann also leicht die Angaben zum Standardfehler ohne Endlichkeitskorrektur nachvollziehen. Da in der Tabelle neben  $n$  auch der entsprechende Umfang  $N$  in der Grundgesamtheit angegeben ist, lässt sich auch der korrekte relative Standardfehler *mit* Endlichkeitskorrektur EK (dem finite multiplier)

$$\sqrt{\frac{N-n}{N-1}} \approx \sqrt{1 - \frac{n}{N}}$$

berechnen. Die EK ist jedoch bei einem kleinen Auswahlsatz  $n/N$  kaum

kleiner als 1, so dass sich  $\hat{\sigma}_{\bar{x}} = \hat{\sigma}_{\bar{x}}^{(o)} \sqrt{(N-n)/(N-1)}$  auch kaum von  $\hat{\sigma}_{\bar{x}}^{(o)}$  unterscheidet. Das zeigt sich auch an der folgenden ergänzenden Tabelle zu Tab. 6.

**Tabelle 7**

Relativer Stichprobenfehler mit Endlichkeitskorrektur in Prozent bei Praxisaufwendungen 2008

Gruppe	$N_k$ bzw. $N$	$n_k$ bzw. $n$	Endl.korr. (EK)	relativer Stichprobenfehler*	
				ohne EK	mit EK
Alle Praxen	95.030	4.172	0,9778	2,449	2,395
Allgemeinmedizin	37.865	319	0,9958 $\approx$ 1	3,816	3,800
Gynäkologie	7.149	415	0,9706	2,756	2,675
Psychotherapie	16.033	447	0,9860 $\approx$ 1	3,187	3,143

\* in vH

Wie man sieht, ändert die Berücksichtigung der Endlichkeitskorrektur nicht sehr viel am Ergebnis der Tab. 6. Die gefundenen Stichprobenfehler sind insgesamt größer als bei der KSE, wo sie gem. Tab. 5 nur um ein Prozent liegen. Diese mögliche Ursache ist nicht belegbar, da im Falle der KSE die veröffentlichten Details zu  $\hat{\sigma}_{\bar{x}_k}$  viel spärlicher sind als beim ZiPP. Der Stichprobenumfang ist in beiden Erhebungen ähnlich,<sup>54</sup> und es ist auch nicht anzunehmen, dass die geschätzten Varianzen der Gesamtaufwendungen in der KSE sehr viel kleiner sind als im ZiPP.

<sup>54</sup> Der faktisch erzielte Auswahlsatz ist bei der KSE für Arztpraxen nach Tab. 2 auch nur 4,59%. Dass in der KSE im Unterschied zum ZiPP auch reine Privatpraxen betrachtet werden dürfte kaum den erheblich geringeren Stichprobenfehler bei der KSE erklären.

Beim Vergleich der relativen Stichprobenfehler ist allerdings zu berücksichtigen, dass Destatis diesen für die Aufwendungen je Praxisinhaber ausweist. Hierbei wird davon ausgegangen, dass sich Einnahmen, Aufwendungen und Jahresüberschuss der Praxis auf die Inhaber gleich verteilen. Das ZiPP weist den relativen Stichprobenfehler auf der Basis der Aufwendungen je Praxis aus. Nach diesem Vorgehen muss der von Destatis ausgewiesene relative Stichprobenfehler kleiner sein, als der vom ZiPP ausgewiesene. Dieser Unterschied kann also nicht als schwächere Repräsentativität interpretiert werden.

### **3. Einzelheiten zum Merkmalskatalog**

Für die Qualität einer Befragung sind „Repräsentativität“ und die Vermeidung (bzw. Quantifizierung) von stichprobenbedingten und nicht-stichprobenbedingten Fehlern wichtige Faktoren, was aber eine Erhebung letztlich leisten kann ist weitgehend bestimmt von der Art und Differenzierung der erfragten Merkmale. Es sollen deshalb in diesem Kapitel die Merkmalskataloge der beiden Erhebungen verglichen werden. Dies kann weitgehend in Form von Übersichten erfolgen.

#### **3.1 Grobübersicht zu den Merkmalskatalogen der beiden Befragungen**

Wir versuchen zunächst einen Überblick zu geben und gehen dann in den Abschnitten 3.2 und 3.3 mehr ins Detail, was die personelle, räumliche und technische Ausstattung der Praxen betrifft, und was die finanziellen Daten (Erträge, Kosten, Vermögen etc.) betrifft.



## Übersicht 4

Übersicht zu den Merkmalskatalogen der beiden Befragungen

	KSE (StBA)	ZiPP (ZI)
Angaben zur Praxis	Grobe Differenzierung nach der Praxisform (Einzelpraxis, Gemeinschaftspraxis und sonstige Praxisform), Facharztbezeichnung der Praxisinhaber, keine Angaben zur Unternehmensdemographie	Detaillierte Erfassung struktureller und personeller (die einzelnen Ärzte betreffend) Merkmale der Arztpraxis (Organisationsform etc.); Daten zur Unternehmensdemographie (Gründung, Organisationswechsel etc.)
Umfang ärztl. Tätigkeit	keine entsprechende Fragen	Angaben zur Zahl und Tätigkeit der Praxispartner, praxisexterne Leistungserbringung, Arbeitszeit des ärztlichen Personals
Einnahmen	Einnahmen aus selbständiger ärztl. Tätigkeit nach GKV und PKV-Einnahmen.	Vertrags- und privatärztliche Umsätze; detaillierte Gliederung vgl. Übers. 7.
Kosten (Aufwendungen)	Personal, Sachmittel <sup>a)</sup> , Mieten/ Leasing <sup>b)</sup> , Abschreibungen, Fremdkapitalzinsen, kein Testat des Steuerberaters bei Fragen nach finanziellen Variablen	Betriebsausgaben nach Kostenarten (wie bei KSE), auch Vermögen u. Finanzierung; bei allen finanziellen Angaben Beantwortung mit Testat durch den Steuerberater
Tätige Personen	Tätige Personen zum Stichtag 30.9.2007; Untergliederung vgl. Übers. 6	Wie KSE, zusätzlich sehr detaillierte Angaben zu Art und Umfang der Tätigkeit der Ärzte und auch des nichtärztl. Personals <sup>c)</sup>
Versorgungsleistungen	keine entsprechende Fragen	Erhebung vieler versorgungsrelevanter Merkmale wie Art und Umfang ärztl. Arbeit; Öffnungszeiten, Behandlungsfälle
Größe/Ausstattung	keine entsprechende Fragen	Fragen zur Praxisgröße, personellen, räuml. u. technischen Praxisausstattung
Anlagevermögen	keine entsprechende Fragen	Anschaffungs- u. Buchwert d. Anlagevermö., Anteil d. Eigenfinanzierung
Meinungsfragen	keine entsprechende Fragen	Bewertung der aktuellen Situation bzgl. Arbeitszeit, Investitionsbedarf, Personalausstatt. Situation als Vertragsarzt

a) Material, fremde Laborarbeiten

b) Bei einer Praxis im eigenen Haus wird der Mietwert bzw. die user costs geschätzt.

c) Arbeitszeiten wurden im ZiPP auch beim nichtärztlichen Personal erhoben.

Ein erheblicher Unterschied zwischen der KSE und dem ZiPP ist die Betonung von Organisationsmerkmalen der Praxen im ZiPP, wie

1. Gründungsjahr (Gründung bzw. Übernahme) und Jahr eines Organisationswechsels (Umwandlung der Organisationsform), Beitritt/Ausscheiden von Partnern bzw. Ärzten, (räumliche) Verlegung
2. Organisationsform und
3. Besondere Versorgungsformen (neu geschaffene Möglichkeiten des „selektiven Kontrahierens“) und sonstige Praxismerkmale, wie Teilnahme am Strukturvertrag, Beteiligung an einem Praxisnetz, Zertifizierung, Qualitätsmanagement.

Mit diesen Merkmalen, die im besonderem Maße eine Panelerhebung rechtfertigen (Organisationsformen sind im Fluss) erhält man zusammen mit Daten zum Leistungsspektrum einer Praxis die Möglichkeit einer sehr differenzierten Praxistypologie (und damit von Determinanten Kosten), was in dieser Detailliertheit bei der KSE nicht darzustellen ist.

Einnahmen sind in der KSE definiert als Einnahmen vor Abzug der Verwaltungskosten der

Kassenärztlichen Vereinigungen und vor Abzug der Gebühren für privatärztliche Verrechnungsstellen.<sup>55</sup> Zur Untergliederung der Einnahmen vgl. Übers. 7.

### 3.2 Struktur der Praxen, Versorgungsleistungen

In diesem Abschnitt geht es um Größe, Rechtsform und ähnliche allgemeine Kennzeichen von Arztpraxen, sowie die eng mit der Größe und Versorgungsleistung einer Praxis zusammenhängende personelle, räumliche und technische Ausstattung der Praxis.

#### Übersicht 5

Praxistyp, Unternehmensdemographie und Versorgungsdaten

	KSE (StBA)	ZiPP (ZI)
Praxisform	Differenzierung nach - Einzelpraxis - Gemeinschaftspraxis (BAG) - Sonstige Praxisform (z.B. Praxisgemeinschaft, MVZ)	Detaillierte Angaben zur Organisationsform wie Einzelpraxis, Kooperationsformen (BAG auch überörtliche und MVZ), Praxisgemeinschaft; Angaben zu den Praxispartnern u. praxisexterne Leistungserbringung (z.B. Belegarzt in OP-Zentren usw.)
Versorgungsleistungen	praktisch keine Angaben zur Versorgungsleistung einer Praxis erfragt	u.a. Teilnahme an "besonderen Versorgungsformen" <sup>a)</sup> , Qualitätsmanagement (Zertifikation), Größe (Radius in km) des Versorgungsgeb.; GKV-Behandlungsfälle nach Quartalen, Anteil der GKV-Patienten an allen Pat.), Arzt-Patienten-Kontakte pro GKV-Fall
Unternehmensdemographie	keine Angaben in diesem Bereich erfragt	Gründungsjahr, verschied. Arten des Organisationswechsels (Ein- und Austritt von Partnern, Anstellung v. Ärzten, Jobsharing, Umwandlung d. Organisationsform, Änderung des Zulassungsumfangs, Verlegung und Übernahme einer Praxis etc.)
Facharztbezeichnung	nur beim Praxisinhaber und Differenzierung nur nach 12 ärztlichen Fachbereichen	Differenzierung nach 36 ärztlichen Fachbereichen (auch Unterscheidung Fach[arzt]-gebietsbezeichnung und Zulassungsfachgebiet)
Arbeits- und Betriebszeiten	Keine Angaben	Differenzierte Erfassung der Arbeitszeiten und deren Aufteilung <sup>b)</sup> bei den Ärzten; Praxisöffnungszeiten (Betriebszeiten) pro Woche und Jahr

a) Strukturverträge, Modellvorhaben, hausarztzentrierte Versorgung, DMP usw.

b) Arbeit mit Patienten, andere ärztliche Tätigkeiten (Gutachten etc.), Managementtätigkeit, Abwesenheitszeiten; Arbeitszeiten wurden auch bei nichtärztlichem Personal erhoben

Wie Übersicht 5 zeigt, bietet das ZiPP erheblich mehr Details hinsichtlich der Praxistypologie (Rechtsform, Anzahl und fachliche Ausrichtung der Ärzte usw.) und Angaben zu Versorgungsleistungen (Leistungsangeboten) der Praxen sowie zur Unternehmensdemographie, die in der amtlichen KSE fast vollständig fehlen. Das ZiPP erlaubt eine erheblich größere Differenzierung nach Facharztbezeichnungen, während in der KSE die Fachrichtung als offene Frage erfragt wird und nur eine Differenzierung nach 12 ärztlichen Fachbereichen der Praxisinhaber erlaubt. Bemerkenswert ist schließlich, dass im ZiPP detailliert Angaben zur räumlichen und techni-

<sup>55</sup> In den allgemein zugänglichen Informationen zur KSE finden sich leider kaum Tabellen mit Differenzierungen der Einnahmen nach GKV- und PKV-Einnahmen obgleich danach gefragt wurde. Angaben über Einnahmen dienen offenbar primär als Bezugsgrundlage für Aufwendungen.

schen Ausstattung der Praxen sowie auch zu den Betriebszeiten erfragt wurden, die auch gerade im Zusammenhang mit den primär interessierenden Kosten stehen, während in der amtlichen KSE derartige Angaben fast vollständig fehlen.

Die laufenden Kosten bestehen überwiegend aus Personalkosten, weshalb Angaben zum Personaleinsatz (Anzahl der Personen, Qualifikation, Arbeitszeit, Art der Tätigkeit) im ZiPP einen großen Raum einnehmen. Beim Leistungsangebot spielt auch die technische Ausstattung eine große Rolle. Hier interessiert nicht nur der Gerätebestand nach Art und Alter und die damit verbunden laufenden Kosten und Investitionen, sondern auch die Einschätzung des künftigen Investitionsbedarfs, die ebenfalls zwar im ZiPP erfragt wurde, nicht aber Gegenstand der KSE war.

Durch im Zeitablauf vergleichbare Angaben zur Größe und Ausstattung der Arztpraxen nach Fachrichtungen und Regionen stellt das ZiPP nicht nur die Kosten- und Einkommensentwicklung im medizinischen Bereich dar, sondern auch die Versorgungssituation für (potentielle) Patienten. Im Vergleich dazu bietet die amtliche KSE eher „nur“ eine Momentaufnahme der Ertrags- und Kostensituation von Arztpraxen, ohne gleichzeitig ähnlich detailliert Bestimmungsfaktoren der Einnahmen und Aufwendungen zu erheben (wie z.B. Behandlungsfälle, Geräteausstattung usw.).

## Übersicht 6

Personelle, räumliche und technische Ausstattung der Praxis

	KSE (StBA)	ZiPP (ZI)
Tätige Personen	Tätige Praxisinhaber, ohne Entgelt mit-helfende Familienangehörige, Lohn- und Gehaltsempfänger (Angestellte Ärzte, technische Assistenten, medizinische Fachangestellte, Azubis u. Sonstiges Personal <sup>a)</sup> )	Anzahl der angestellten Ärzte <sup>b)</sup> und angest. Arzthelferinnen (MFA), Azubis, Krankenschwestern/-pfleger, MTA/MTRA, spez. qualif. Pers., Bürokräfte, Familienangehörige, Sonstiges und externes Personal
Arbeitszeit	keine Fragen nach Arbeitszeiten	Arbeitszeit der Ärzte und des anderen Personals gem. obiger Gliederung, Überstunden u. vertragl. Arbeitszeit <sup>c)</sup>
Praxisfläche, Räume	keine entsprechende Fragen	Fläche der Praxis, spez. d. Funktionseinheiten (auch Anz. d. Zimmer), Vorhandensein best. Funktionseinheiten <sup>d)</sup>
Medizin. techn. Geräte	keine entsprechende Fragen	Anzahl der Geräte auch Art der Beschaffung (Kauf, Miete/Leasing), Anschaffungsjahr u. Restlebensdauer; geschätzter Investitionsbedarf (in €) in Geräte 2010

- a) Gesonderte Darstellung der weiblichen, der in Teilzeit und der sich in Teilzeit befindlichen weiblichen tätigen Personen.
- b) Gesonderte Darstellung ärztl. Weiterbildungsassistenten/innen (WBA).
- c) Wöchentliche Arbeitszeit, auch Arbeitszeit als Belegarzt in OP-Zentren usw.
- d) Art der Geräte bzw. Funktionseinheiten (z.B. Raum für Ultraschall) verschieden je nach ärztl. Fachbereich.

### 3.3 Einnahmen, Ausgaben und Vermögen im Detail

Übersicht 7a stellt den Erhebungsumfang hinsichtlich der Einnahmen (Umsätze) dar, aus denen sich zusammen mit den Kosten der Reinertrag<sup>56</sup> ergibt. In der KSE werden drei Einnahmearten unterschieden; im ZiPP umfasst dieser Bereich jedoch sechs Fragen (die Nummern sind in der Übersicht angegeben), die in dem vom Steuerberater zu testierenden Bereich der

<sup>56</sup> In beiden Erhebungen definiert als Summe der Einnahmen minus Summe der Aufwendungen.

Befragung stehen.<sup>57</sup> Die Einnahme-Aggregate dürften in beiden Erhebungen ähnlich abgegrenzt sein. Einnahmen aus mit Versicherungen (PKV und GKV) abgerechneten Leistungen sind Einnahmen vor Abzug der Verwaltungskosten und Gebühren für Verrechnungsstellen.

Hinsichtlich der Kosten (vgl. Teil b der Übers. 7) ist die Entwicklung der Personalkosten (als wichtigster Kostenfaktor) natürlich besonders relevant. Daneben finden die Abschreibungen besonders Beachtung, weil sie in letzter Zeit zurückgegangen sind<sup>58</sup> und die Erklärung dieses Vorgangs wichtige Fragen aufwirft. Ein Rückgang der Abschreibungen weist auf eine geringere vergangene Investitionstätigkeit hin, die entsprechend eine geringere (aus den Abschreibungen zu finanzierende) Ersatzinvestition (Reinvestition) nach sich zieht. Wie dies zu beurteilen ist hängt vom (naturgemäß nur zu schätzenden) Investitionsbedarf ab, der wiederum eine Funktion des für erforderlich gehaltenen Kapitalstocks ist. Es ist deshalb konsequent, wenn im ZiPP auch nach den tatsächlichen Buchwerten des vorhandenen Kapitalstocks (Anlagevermögens) und der Einschätzung des Investitionsbedarfs gefragt wird. Im Gegensatz zur Höhe der Neu-Investitionen und des Anlagevermögens handelt es sich bei den Angaben zum offenen (nicht befriedigten) Investitionsbedarf um (subjektive) Einschätzungen der befragten Praxisinhaber.

## Übersicht 7

Details zu Fragen nach Einnahmen und Aufwendungen

a) Erträge

	KSE (StBA)	ZiPP (ZI)
Einnahmen	Bruttoeinnahmen aus selbstständiger ärztlicher Tätigkeit aus 1. ambulanter u. stat. Kassenpraxis 2. ambulanter und stat. Privatpraxis 3. sonstiger selbst. ärztl. Tätigk. <sup>a)</sup>	Einnahmen aus ambulanter und beleg-ärztl. GKV-Praxis (B21) und Privatpraxis (B22) <sup>b)</sup> , Sonstige ärztliche (B23) u. nichtärztliche (B24) Tätigkeit <sup>c)</sup> , Sonderbetriebseinnahmen (B26); in Einnahmen insges. (= B25) enthaltene Sonderbetriebseinnahmen

- a) Z.B. Einnahmen aus betriebsärztlicher Tätigkeit, aus nebenamtlicher Krankenhaus-tätigkeit, aus durchgangsarztlicher Tätigkeit, aus selbstständiger Vertretung sowie aus Gutachter-tätigkeit.
- b) Zu PKV-Einnahmen (oder Einnahmen aus Privat-Praxis) gehören auch Abrechnungen zu Lasten von Berufsgenossen-schaften und bei Selbstzahlern abgerechnete Leistungen.
- c) Dazu gehört auch Einnahmen aus der Überlassung von Infrastruktur, Geräten, Flächen u. ä. sowie für die (fremde) Leistungserbringung.

Dass nicht nur eine ständige Beobachtung der laufenden Betriebskosten, sondern auch der Investitionskosten verlangt wird ergibt sich auch aus den eingangs zitierten Rechtsvorschriften, wenn z.B. in § 87, Absatz 2f von der „Entwicklung der für Arztpraxen relevanten Investitions- und Betriebskosten“ die Rede ist.

Die mehr oder weniger detaillierte Aufgliederung der laufenden Aufwendungen von Praxen in den beiden Erhebungen ist unumstritten. Die Kosten umfassen Aufwendungen für Personal, Praxisräume, Material und Labor, Energie, Versicherungen, Kraftfahrzeughaltung usw., sowie Aufwendungen für Leasing und Miete von Geräten und Fremdkapitalzinsen. Sie umfassen z.B. auch Aufwendungen für die Fortbildung, nicht aber für die Alters-, Invaliditäts-, Hinterbliebenen- und Krankenversicherung der Ärzte, die aus dem Reinertrag zu finanzieren sind. Aufwendungen für Personal dürften in beiden Erhebungen ähnlich abgegrenzt sein und neben Bar- auch Sachbezüge sowie Steuern und Sozialversicherungsbeiträge der Arbeitnehmer enthalten. Die Arbeitgeberanteile zur Sozialversicherung gehören nicht zu den Bruttolöhnen und -gehältern und sind eine gesonderte Position, zusammen mit übrigen Sozialaufwendungen,

<sup>57</sup> Das gilt auch für die übrigen im Bereich dieses Abschnitts behandelten Fragen (Merkmale) der Erhebung.

<sup>58</sup> vgl. Jahresbericht des ZI zum ZiPP 2010, S. 16

die auf tariflicher, betriebs- und branchenüblicher Grundlage bzw. freiwillig gewährt werden.<sup>59</sup>

## Übersicht 7

### b) Kosten

	KSE (StBA)	ZiPP (ZI)
Personal- aufwen- dungen	1. Bruttolöhne und -gehälter 2. gesetzl. u. übr. Sozialaufwend. 3. Honorare <sup>a)</sup>	Aufwendungen f. Personal (differenziert nach internes (B11), und externes Personal (B12); darin auch ärztl. Praxisvertretung)
Material, Labor	Aufwend. f. Material u. Arbeit bei eigenen u. fremden Laborarbeiten	Material, Labor (B13), ähnlich wie nebenstehend abgegrenzt <sup>b)</sup>
Mieten/ Leasing für Räume <sup>c)</sup>	effektiv 2007 zu zahlende Mieten f. d. Praxisräume (inkl. Garagen); Nebenkosten als gesonderte Position	Miete f. Praxisräume bzw. Mietwert (B14), Aufwendungen für Nebenkosten (Heizung, Strom, Gas, Wasser) (B15)
Geräte	Miete/Leasing Apparate, EDV etc.	Leasing/Miete v. Geräten (B19)
Versiche- rungen etc.	Beiträge für Versicherungen <sup>d)</sup> Beitr. zu Berufsorganisationen <sup>e)</sup>	Versicherungen, Beiträge u. Gebühren (B16); Definition wie bei KSE
Kfz-Kosten	praxisbed. Kfz-Aufwendungen <sup>f)</sup>	anteilige (praxisbedingte) Kfz Kosten (B17)
Abschrei- bungen <sup>g)</sup>	steuerlich absetzbare Abschreibungen auf die Praxiseinrichtung	Abschreibungen (B18) auf Praxiseinrichtung, immaterieller Praxiswert (B18.1 u. 2)
Geringwer- tige Güter <sup>g)</sup>	absetzbare Wirtschaftsgüter unter 410 € Anschaffungskosten	geringwertige Güter GWG (B18.3) gem. § 6 EStG; Def. wie bei KSE
Fremdkap.	Fremdkapitalzinsen <sup>h)</sup>	Fremdkapitalzinsen (B110)
Fortbildung		gesondert: Aufwend. f. Fortbild. (B111)
Sonstige Aufwend.	Sammelposition für übrige Aufwendungen <sup>i)</sup>	(B112) Sonstige: Wartung, Reparatur, Qualitätssicherung, EDV, Beratung etc.
Sonderbe- triebsausg.		in Aufwendungen enthaltene Sonderbetriebsausgaben (B114)

a) für Assistenz und Praxisvertretung

b) Berücksichtigt werden auch Verlustbeteiligungen aus Laborgemeinschaften.

c) Mieten für privat genutzte Räume nicht einbezogen; Schätzung des Mietwerts wenn Praxis im eigenen Haus betrieben wird mit Kosten des eigenen Grundstücks und Gebäudes (soweit sie auf die Praxis entfallen), z.B. Hypotheken- und Grundschuldzinsen, Gebäudeabschreibungen, Instandhaltungskosten.

d) Berufshaftpflicht- und Praxisversicherung, Feuer-, Diebstahlversicherung usw. ohne Kraftfahrzeugversicherung

e) ohne KV- bzw. KZV-Verwaltungskosten

f) Anteilige Kfz-Steuer, Haftpflicht-, Kasko-, Insassenunfall- und Rechtsschutzversicherung, Abschreibungen, Reparatur- und Instandhaltungskosten, Kraftstoff- und sonstige laufende Betriebskosten; im Falle von Leasing-Verträgen für Kraftfahrzeuge gehören auch die anteiligen Mietraten (Leasing-Raten) zu den praxisbedingten Kraftfahrzeugaufwendungen.

g) Abnutzbare, bewegliche und selbstständig nutzbare Güter, wie Instrumente, Apparate etc., inkl. Labor mit einem Anschaffungswert von mehr als 410 €, die voll als Praxisausgaben absetzbar sind. Sonderabschreibungen sind eingeschlossen.

h) Zinsen für das der Praxis zur Verfügung gestellte Fremdkapital (z.B. zur Anschaffung von Einrichtungsgegenständen oder zur Kapitalbeschaffung bei der Praxisaufnahme). Nicht eingeschlossen sind Zinsen für Darlehen, die für bauliche Maßnahmen aufgenommen worden sind.

i) Aufwendungen für wissenschaftliche Kongresse, Fortbildungskurse, Fachliteratur und dgl., betriebliche Steuern, Büromaterial, Post- und Telefongebühren, Berufsbekleidung, Gebühren an privatärztliche Verrechnungsstellen, Steuerberatung, Buchführung und Rechtsberatung, Reisekosten, Reparaturen, EDV-Kosten, Kostenerstattung an Belegkrankenhäuser. Nicht enthalten sind die als außerordentlich und betriebsfremd anzusehenden Aufwendungen und die Einkommensteuer.

59 Beispielsweise Aufwendungen für die betriebliche Altersversorgung, Beiträge zur Aus- und Fortbildung, Zuschüsse im Krankheitsfall, Zuschüsse für Verpflegung und Kleidung, Trennungschadigungen und Umzugskostenvergütungen.

## Übersicht 7

### c) Vermögen, Investitionen

	KSE	ZiPP (ZI)
Vermögen	Keine Fragen hierzu	Anschaffungswert (B31) u. Buchwert (B34) d. Anlagevermög., Übernahmementgelt (B32) <sup>a)</sup> , nachträglich angeschafftes Vermögen (B33)
Investition		Neuinvestitionen 2006, 2007, 2008 (B35)
Eigenfinanzierung		Anteil der Finanz. aus Eigenmitteln bei Investitionen in Ausstattung, Geräte, Goodwill (B4)

a) differenziert nach materiellem und immateriellem Praxiswert

## 4. Besondere methodische Aspekte

In diesem Kapitel wird auf einige bereits angesprochene Probleme weiter eingegangen. Das betrifft lediglich das ZI Praxispanel (ZiPP), nicht aber die amtliche Kostenstrukturerhebung. Betrachtet werden vor allem die Vor- und Nachteile von Panelerhebungen (Abschn. 4.1) und die „Standardisierung“ von Erträgen, insbesondere des Reinertrags einer Arztpraxis.<sup>60</sup>

### 4.1 Bemerkungen zur Panelbefragung

Echte Längsschnittsbetrachtungen im Sinne von individualisierten Verlaufserhebungen sind sehr aufwändig (und entsprechend selten), weil sie eine Identifikation der Einheit verlangen (um festzustellen, dass es sich bei der Einheit  $i$  in der Erhebung in  $t$  um die gleiche Einheit handelt, wie  $j$  in  $t-1$ ) und sie sind deshalb i. d. R. nur geboten, wenn es gilt,

1. Entwicklungen im Zeitablauf frei von Strukturveränderungen darzustellen oder
2. Verweildauern festzustellen und Kohorteneffekte herauszuarbeiten.<sup>61</sup>

Beim ZiPP spielt vor allem die erste Überlegung eine Rolle. Ein Panel<sup>62</sup> dient der Feststellung „echter“ Veränderungen im Unterschied zur Zeitreihenbetrachtung (repeated observations), bei der sich die Struktur der erfassten (Teil-)Gesamtheit zwischen den Erhebungszeitpunkten durch Zu- und Abgänge laufend ändern kann (so dass in den Ergebnissen „echte“ und strukturell bedingte Veränderungen schwer oder gar nicht zu unterscheiden sind). Bei dieser Zielsetzung eines Panels ist es geboten, vorzeitig ausscheidende Einheiten nicht mehr zu berücksichtigen, die Auswertung also zu beschränken auf alle die Einheiten, die im gesamten Untersuchungszeitraum in der Stichprobe vorhanden waren.

Anders verhält es sich bei der zweiten oben genannten Fragestellung (Kohorteneffekte) und den damit vielleicht (gerade Medizinern) bekannteren Anwendungen von Längsschnittsbetrachtungen, wenn es nämlich gilt, die Zeit selbst als Merkmal zu betrachten, z. B.

- die Zeitpunkte von Übergängen<sup>63</sup> zwischen Zuständen und damit der Verweildauer in bestimmten Zuständen festzustellen, wie z.B. die Dauer (und Wahrscheinlichkeit)

60 Der in der Öffentlichkeit besonders beachtete Reinertrag stellt nicht den betriebswirtschaftlichen Gewinn der Praxis dar, da er u.a. noch die Aufwendungen für eine Praxisübernahme oder für die Alters-, Invaliditäts-, Hinterbliebenen- und Krankenversicherung der Praxisinhaber und der Familienangehörigen, und auch Beiträge zu Versorgungseinrichtungen der Ärzte für das Jahr 2007 enthält.

61 Im früher üblichen Sprachgebrauch wurde von Längsschnittsanalysen nur im Sinne einer echten Verlaufsanalyse gesprochen, die meist in Form von Kohortenanalyse durchgeführt wurden. In der Ökonometrie wird oft von „Längsschnitt“ im Sinne einer Zeitreihe gesprochen, also auch dann wenn sich die Struktur der befragten Anwesenden permanent ändert.

62 Mehr hierzu vgl. Peter von der Lippe, Verlaufsanalysen (Panelerhebungen) in der Statistik: Warum und wie? DISKUSSIONSBEITRAG aus der Fakultät für WIRTSCHAFTSWISSENSCHAFTEN der UNIVERSITÄT DUISBURG – ESSEN, Campus Essen, Nr. 186, Januar 2011, als Download unter [von-der-lippe.org/publikationen4.php](http://von-der-lippe.org/publikationen4.php).

63 Man kann allerdings oft auch aus Querschnittsdaten Übergangswahrscheinlichkeiten (transition probabilities) schätzen und damit unechte Längsschnittsanalysen herleiten. Ein Beispiel hierfür sind Sterbetafeln (Berechnung von Lebenserwartungen) mit aus Querschnitten ermittelten (einjährigen) Sterbewahrscheinlichkeiten.

- des Überlebens einer Operation als Indikator eines Therapieerfolgs,<sup>64</sup>
- drop outs zu erfassen (z.B. Studienabbrecher, Studienfachwechsler)<sup>65</sup> oder
- zwischen Alters-, Perioden- und Kohorteneffekten zu differenzieren.

Bei derartigen Anwendungen ist es natürlich nicht geboten, vorzeitig ausscheidende Einheiten in der Erhebung insgesamt (also von Anfang des Intervalls einer Auswertung an) nicht zu berücksichtigen, weil ja gerade der Zeitpunkt des Ausscheidens von Interesse ist. Der Vorteil dieser Anwendungen der Panelmethodik<sup>66</sup> ist, dass (nur) so *individuelle* (oder „intra-individuelle“) Verläufe sichtbar gemacht werden können, nicht nur Veränderungen auf der Aggregatebene (also *durchschnittliche* Veränderungen).

Man unterscheidet die Generation (Kohorte)  $g$ , das Alter  $x$  und die Periode  $t$ , wobei die drei Größen untereinander zusammenhängen  $g + x = t$ . Dass man in  $t = 2012$  genauso wie im Jahr  $t = 1970$  mit zunehmendem Alter ein höheres Morbiditätsrisiko hat, ist ein Alterseffekt. Es kann aber auch sein, dass sich die subjektiv wahrgenommenen Gesundheit, Lebensweise, Ernährungsverhalten und Inanspruchnahme von Gesundheitsleistungen der heute Sechzigjährigen (also des Geburtsjahres 1952) signifikant unterscheidet von dem der Sechzigjährigen vor fünf Jahren (also der Kohorte  $g = 1947$ ). In diesem Fall spricht man von einem Kohorten- oder Generationeneffekt.<sup>67</sup>

Für derartige Betrachtungen sind wiederholte Querschnitterhebungen nicht sachgerecht, weil bei ihnen immer nur die in  $t$  „Anwesenden“ betrachtet werden, und nicht jeweils eine Kohorte aus gleichen Einheiten über längere Zeiträume.<sup>68</sup> Panel als wesentlich aufwändigere Erhebungen haben jedoch mit typischen Schwierigkeiten zu kämpfen, wie

- *Panelmortalität (panel attrition)*<sup>69</sup> die dazu führt, dass die Zahl der permanent im Panel befindlichen Einheiten stark zusammenschrumpfen kann, weil es schwer wird, über einen langen Zeitraum genügend „matched pairs“ und damit ein ausreichend großes „balanced panel“ zu haben, und damit zusammenhängend
- ein unit nonresponse Problem im Sinne von „Nicht-mehr-Beantwortung“ (im Unterschied zur Nichtbeantwortung, die schon in der ersten Befragungswelle auftreten kann, können hier Angaben der betreffenden Einheit aus früheren Wellen evtl. eine sog. „Imputationen“ erleichtern),<sup>70</sup> und schließlich
- *panel conditioning* (Paneffekt), was z.B. bei wiederholten Meinungsbefragungen der gleichen Panelteilnehmer (panelists) über einen längeren Zeitraum eine Änderung eintreten kann, für die die Befragung selbst die Ursache sein kann (z.B. Neubewertung aufgrund der Befragung oder Anpassung an Meinungen, die als gewünscht vermutet werden).<sup>71</sup>

Panelmortalität und damit verbundene Antwortausfälle wären dann kein Problem, wenn sie zufällig verteilt wären. Häufig sind die Ausfälle aber systematisch (korreliert mit Merkmalen der

64 Man spricht bei der Dauer von Arbeitslosigkeit, Krankheit, Schulbesuch etc. auch von „Episoden“ [spells].

65 Die Gegenstände hängen untereinander zusammen. So ist z.B. die zuverlässige Feststellung von Verweildauern nicht möglich, wenn es nicht gelingt „drop outs“ zu erkennen.

66 Beliebte sind auch als Panelerhebungen dieser Art Befragungen nach Meinungen der gleichen Person zum gleichen Thema im Zeitablauf oder zur Veränderung des Sozialstatus (Analysen des sozialen Wandels).

67 Aus solchen Gründen gibt es auch entsprechende Längsschnittanalysen in England (English Longitudinal Study of Ageing ELSA), den USA (Health and Retirement Study HRS) und auf EU-Ebene (SHARE = Survey of Health, Ageing and Retirement in Europe). Kohorteneffekte herauszuarbeiten verlangt natürlich insgesamt einen sehr langen Erhebungszeitraum. Im obigen Beispiel ( $g = 1947$  und  $g = 1952$ ) dürfte der zeitliche Unterschied der Kohorten wohl noch zu gering sein.

68 Es gab früher einmal in der amtlichen Statistik in Gestalt der Statistik der Aktiengesellschaften eine Erhebung, die zwar jeweils eine Querschnittserhebung war, bei der man sich aber in den Veröffentlichungen auf die Einheiten beschränkte, die in jeweils  $q = 2$  aufeinanderfolgenden Querschnittserhebungen (nicht aber über einen längeren Zeitraum in allen  $q > 2$  Erhebungen) gleichermaßen in der Beobachtungsgesamtheit vertreten waren.

69 Das Problem ist gut beschrieben mit „disproportionate dropout“.

70 Der Umstand, dass eine Einheit früher einmal teilgenommen hat erleichtert i.d.R. das Ersetzen fehlender Angaben durch Schätzung (d.h. Imputation). Im Gegensatz dazu sind bei der üblichen non-response Problematik einer Einmalbefragung i.d.R. überhaupt keine Erkenntnisse über die Antwortverweigerer zu erhalten.

71 Derartige Effekte von wave specific responses sind wohl im Falle des ZiPP nicht zu erwarten.

Einheiten), weshalb man auch von „selektiver“ panel attrition spricht, und die oben in Abschn. 2.4 beschriebene einfache Hochrechnung nicht voll befriedigen kann.

Im Folgenden können wir nur das Problem, systematischer Ausfälle kurz einführen, mögliche Methoden aber nicht weiter erläutern. Solche Ausfälle zu berücksichtigen ist schwierig, weil i.d.R. über die nichtteilnehmenden Einheiten (Subskript  $m = \text{missing}$ ) keine Informationen vorliegen, außer denen des sampling frames (hier das BAR), oder evtl. ausser Informationen aus einer früheren Befragungswelle, und weil die Gruppe meist anders strukturiert sein wird als die Gruppe der teilnehmenden Einheiten (Subskript  $r = \text{responding}$ ). Der Stichprobenumfang  $n$  setzt sich zusammen aus<sup>72</sup>  $n = r + m$  Einheiten und die Grundgesamtheit aus  $N = R + M$  Einheiten. Entsprechend gilt für

$$\text{Mittelwerte in der Stichprobe } \bar{y}_n = \frac{r}{n} \bar{y}_r + \frac{m}{n} \bar{y}_m$$

und es gilt  $\bar{y}_n$  zu schätzen obgleich nur  $\bar{y}_r$ , nicht aber  $\bar{y}_m$  bekannt ist. Für die Punktschätzung (von  $\mu$  [oder  $\mu_y$ ] aufgrund von  $\bar{y}$ ) und für die „Hochrechnung“ von Merkmalssummen (von  $\Sigma y_j$  [ $j = 1, \dots, N$ ] in der Grundgesamtheit aufgrund von  $\Sigma y_i$  [ $i = 1, \dots, r$ ] in der Stichprobe) sind drei Parameter relevant, nämlich die

- die Nonresponse-Quote  $m/n$  (zusammen mit dem Auswahlatz  $n/N$ ), die
- Unterschiedlichkeit der Mittelwerte  $\bar{y}_r$  und  $\bar{y}_m$  und die
- Kovarianz zwischen dem Response-Indikator<sup>73</sup>  $\rho_i$  und der zur Diskussion stehenden Variable  $y_i$

die für den „non-response bias“ (ein „bias“ ist ein systematischer Fehler) verantwortlich sind. Die üblichen – aber nicht unproblematischen – Methoden im Umgang mit Nonresponse sind

- *ignorieren*
- *redressment* der Stichprobe (mit geeigneten Gewichtungsverfahren) und
- *imputation* (Schätzung eines oder mehrerer Werte für den fehlenden Wert [eine naheliegende Option ist z.B. den Mittelwert der Teilgesamtheit zu der die missing unit gehört als Schätzung der missing observation einzusetzen]).

Bei der sog. einfachen Imputation wird für einen von der nichtantworteten Einheit  $i$  gelieferten Wert  $y_i$  ein Schätzwert  $\hat{y}_i$  eingesetzt, der sich aus der (mit Daten der – auch auf  $y$  – antwortenden Einheiten geschätzten) Regressionsfunktion  $\hat{y} = f(x_1, \dots, x_K)$  ergibt (dabei ist natürlich vorauszusetzen, dass es bei Einheit  $i$  zwar kein Wert für  $y$  gibt, wohl aber Angaben für die Regressoren  $x_1, \dots, x_K$  existieren). Auf weitere Hinweise zu den Methoden bei Antwortausfall kann hier nicht weiter eingegangen werden.<sup>74</sup> Welche Methode angemessen ist hängt von Zufälligkeit/Nichtzufälligkeit (oder Korreliertheit) der Nichtbeantwortung (bzw. des Ausfalls einer Einheit im Panel) ab.

## 4.2 Standardisierung der Einnahmen als Besonderheit des ZiPP

Wie bereits erwähnt sind für bestimmte Betrachtungen Umrechnungen der tatsächlichen Einnahmen einer Praxis in solche (notwendig fiktive) Einnahmen einer Praxis gewünscht, die ausschließlich GKV-Patienten hat und Leistungen eine Norm- oder „Standard“-Arbeitszeit nicht über- oder unterschreitet. Bei der Auswertung des ZiPP wird – im Unterschied zur KSE des StBA – eine „Standardisierung“, d.h. Umrechnung der Einnahmen auf ein Niveau, das sich bei einer Standardpraxis (Anteil  $p$  von Privatpatienten in Höhe von  $p = 0$  und Arbeitszeit  $A = A^* = 140.148$  Minuten) ergeben hätte.<sup>75</sup> Für die Standardisierung bezüglich der Arbeitszeit (also

<sup>72</sup> Die entsprechende Betrachtung wird ungleich komplizierter, wenn man auch zwischen non-response und einem withdrawal of a panelist (also dem Problem der Panelmortalität, dass eine bekannte Einheit, die früher teilgenommen hat, bei dieser Welle nicht mehr antwortet) unterscheiden will.

<sup>73</sup> Eine Indikatorvariable ist eine 0-1 Variable. Es gilt  $\rho = 1$  wenn Teilnahme und  $\rho = 0$  wenn keine Teilnahme.

<sup>74</sup> Vgl. hierzu die bereits zitierte Arbeit „Das PHP: eine Erhebung zu Vermögen...“ der Deutschen Bundesbank.

<sup>75</sup> Die Kritik an der vorgefundenen Methode (der sog.  $\eta$ -Methode) und meine Weiterentwicklung (zur  $\lambda$ -Methode) ist ausführlich beschrieben in Peter von der Lippe, Standardisierung der Einnahmen einer Arztpraxis Methoden der Honorarumrechnung einer



für die Berücksichtigung des [zeitlichen] Auslastungsgrads) scheint die Division der tatsächlichen Einnahmen  $E$  durch  $\kappa = A/A^*$  unproblematisch zu sein, so dass bei Mehrarbeit ( $\kappa > 1$ ) einer Praxis  $i$  die standardisierten Einnahmen  $\hat{E}_i = E_i/\kappa_i$  geringer und bei Unterauslastung ( $\kappa < 1$ ) größer sind als die tatsächlichen Einnahmen  $E_i$ . Diese Betrachtungsweise impliziert einen linearen Zusammenhang (eine Proportionalität) zwischen  $E_i$  und  $A_i$  (insbesondere auch  $E_i = 0$  wenn  $A_i = 0$  ist oder wenn  $A_j > A_i$ , dann auch  $E_j > E_i$ ), was plausibel sein dürfte.

Die Standardisierung bezüglich der Struktur der Patienten scheint dagegen eher problematisch zu sein, insbesondere was die vorgefundene Methode einer Division der Einnahmen

$$\text{durch } \eta_i = \frac{u_i}{p_i}$$

betrifft, wobei  $u_i = E_{P_i}/E_i$  als Anteil der PKV-Umsätze<sup>76</sup>  $E_{P_i}$  an allen Umsätzen  $E_i$  definiert ist und  $p_i = N_{P_i}/N$  als Anteil der PKV-Patienten (Anzahl  $N_{P_i}$ ) an allen  $N$  Patienten.

Aus diesen Definitionen ergibt sich

$$(3) \quad e_i = e_{G_i}(1-p_i) + p_i e_{P_i}.$$

Mit dem Aufschlagsfaktor  $\alpha_i = e_{P_i}/e_{G_i}$  erhält man

$$(4) \quad E_i = N_i e_i = N_i e_{G_i}[(1-p_i) + p_i \alpha_i] = N_i e_{G_i}[1 + (\alpha_i - 1)p_i] = N_i e_{G_i} \lambda_i$$

wobei  $N_i e_{G_i}$  der gesuchte Umsatz einer gleichgroßen (gleiche Anzahl  $N$  der Patienten) Normpraxis wäre (d.h. einer Praxis  $i$ , die nur GKV-Patienten hat, also mit  $p_i = 0$ ). Die Standardisierung kann also erfolgen, indem man den Gesamtumsatz (Gesamteinnahmen) einer Praxis  $E_i$  durch

$$(4a) \quad \lambda_i = 1 + (\alpha_i - 1)p_i$$

statt durch  $\eta_i = e_{P_i}/e_i$  dividiert. Wie man sieht, ist  $\lambda_i$  umso größer, je größer bei gegebenem  $\alpha_i$  der Privatpatientenanteil  $p_i$  ist. Die Größe

$$\eta_i = \frac{\alpha_i}{1 + (\alpha_i - 1)p_i} = \frac{\alpha_i}{\lambda_i}$$

nimmt dagegen bei gegebenem  $\alpha_i$  ab, statt zu, wenn  $p_i$  steigt. Paradox ist auch die Situation  $p_i = 1$  und damit auch  $u_i = 1$ , was zu  $\eta_i = 1$  führt, also bedeuten würde, dass keine Standardisierung erforderlich ist. Dabei ist eine solche Praxis das Gegenteil der gewünschten reinen GKV-Praxis, weil alle ihrer Patienten PKV-Patienten sind. Nimmt man das  $\alpha$ -fache pro PKV-Patient ein, so müssten die Einnahmen dieser Praxis auch das  $\alpha$ -fache der Normpraxis sein und die Einnahmen durch  $\alpha$  zu dividieren sein. In der Tat ist in diesem Fall ( $p_i = 1$ )  $\lambda_i$  auch gleich  $\alpha_i$ .

Wie man an der Gleichung für  $\lambda_i$  (Gl. 4a) sieht, verlangt die Bestimmung von  $\lambda_i$  Angaben über den Privatpatientenanteil  $p_i$  (und damit den Anteil  $1 - p_i$  der GKV-Patienten), sowie über die entsprechenden Einnahmen, so dass die Pro-Patient erzielten Umsätze  $e_P$  und  $e_G$  bestimmen lassen, aus denen sich  $\alpha_i = e_{P_i}/e_{G_i}$  errechnet. Die entsprechenden Angaben waren nicht in allen Fällen vorhanden: die hochgerechnete (gewichtete) Zahl der Praxen 2008 war 4.172 (vgl. Tab. 4), eine Standardisierung (d.h. Bestimmung von  $\lambda_i$ ) war aber nur bei 3.476 (also

Arztpraxis auf Einnahmen einer „Normpraxis“, die ausschließlich EBM-Leistungen (für GKV-Patienten) in Vollzeit erbringt, IBES-DISKUSSIONSBEITRAG, UNIVERSITÄT DUISBURG – ESSEN, Campus Essen, Heft 191 (November 2011); auch als Download unter [von-der-lippe.org/publikationen4.php](http://von-der-lippe.org/publikationen4.php).

76 bzw. der Nicht-GKV-Umsätze

83,3%) Praxen möglich.<sup>77</sup> Es lassen sich also i.d.R., aber nicht immer, mit den ZiPP Daten die Größen  $\lambda_i$  und  $\alpha_i$  bestimmen, so dass kein Rückgriff auf externe Angaben (z.B. über den Aufschlagsfaktor gegliedert nach Arztgruppen) für eine Standardisierung mit der „ $\lambda$ -Methode“ erforderlich ist.<sup>78</sup> Will man allerdings die praxisindividuellen Größen  $\alpha_i$  zu gruppenspezifischen (z.B. für eine Facharztgruppe) Aufschlagsfaktoren  $\alpha$  durch ein gewogenes Mittel aggregieren, so bieten sich verschiedene Möglichkeiten für eine Gewichtung an. Das könnte von Interesse sein, um die so aus den  $\alpha_i$  der ZiPP Daten bestimmten Aufschlagsfaktoren (als Maß der Mehrvergütung durch die PKV) mit anderen Schätzungen zu vergleichen.

Man kann die beiden Standardisierungen mit  $\lambda_i$  und  $\kappa_i$  nacheinander ausführen, so dass das Gesamtergebnis durch das Produkt der beiden Faktoren  $\lambda_i$  und  $\kappa_i$  bestimmt ist. Im Jahresbericht 2010 zum ZiPP finden sich Angaben zum Verhältnis des standardisierten zum nicht standardisierten Jahresüberschuss. Im Schnitt (auf alle Praxen bezogen) ist der standardisierte Jahresüberschuss nur 78% des nicht standardisierten Jahresüberschusses was einem Wert von  $\lambda\kappa = 1/0,78 = 1,28$  entspricht, wobei, aus Gründen, auf die wir gleich zurückkommen werden,  $\kappa$  wohl mehr als  $\alpha$  zu einem Wert  $\lambda\kappa > 1$  beigetragen haben dürfte. Tab. 8 zeigt einige Ergebnisse für die Größe  $\lambda\kappa$ , die, wie nicht anders zu erwarten war, je nach Arztgruppe sehr unterschiedlich ist.

**Tabelle 8**

Impliziter Faktor  $\lambda \cdot \kappa$  bei einigen standardisierten Jahresüberschüssen (ZiPP)

	standardis. Übersch. in % des nichtstand. Überschuss	impliziter Faktor $\lambda\kappa$
Allgemeinmedizin und Innere Medizin (hausärztlich)	90	1,11
Chirurgie	59	1,69
Neurologie, Nervenheilkunde, Psychiatrie	73	1,37
Orthopädie	54	1,85
Radiologie und Nuklearmedizin <sup>a)</sup>	23	4,35
Psychotherapie <sup>b)</sup>	107	0,93

a) Größter Wert für  $\lambda\kappa$ ; Gründe könnten sein zeitliche Überauslastung oder geringe Vergütung durch GKV relativ zu PKV; zu beachten ist auch, dass die Anzahl der Praxen in der Stichprobe gering ist.

b) Wohl niedrigster Wert für eine Arztgruppe; Gründe evtl. umgekehrt wie bei a (auch was die Anzahl der Praxen betrifft).

Quelle: ZI-Praxis-Panel Jahresbericht 2010, S. 47, Tab. 22

Eine Häufigkeitsverteilung der Lambda-Werte findet sich als Abb. 5 im Jahresbericht 2010 des ZI zum ZI-Praxis-Panel (ZiPP).<sup>79</sup> Danach war bei etwa 29% der befragten Praxen  $\lambda$  kleiner als 1, und bei ca. 43% war  $\lambda$  nur zwischen 1 und 1,3. Wir haben diese Zahlen der Abb. 2 in der folgenden Tabelle 9 wiedergegeben.<sup>80</sup>

Dabei fällt auf, dass die Werte für  $\lambda_i$  sich nicht wesentlich von 1 unterscheiden. Andererseits hängen  $\lambda_i$  und  $\alpha_i$  gem. (4) eng miteinander zusammen, denn aus (4a) folgt

77 Das ergibt sich durch den Vergleich von Tab. 25 und 26 im ZiPP-Jahresbericht 2010 des ZI.

78 Gedacht war dabei an den „Aufschlagsfaktor“, der mir aus einer Arbeit von Wasem et al. bekannt war (und der sich allerdings auch nicht auf Durchschnittseinnahmen je Patient [PKV oder GKV], sondern auf vergleichbare ärztliche Leistungen an den betreffenden Patienten bezieht). Vgl. Anke Walendzik, Stefan Greß, Maral Manougian und Jürgen Wasem, Vergütungsunterschiede im ärztlichen Bereich zwischen PKV und GKV auf Basis des standardisierten Leistungsniveaus der GKV und Modelle der Vergütungsangleichung, DISKUSSIONSBEITRAG aus dem Fachbereich Wirtschaftswissenschaften, UNIVERSITÄT DUISBURG – ESSEN, Campus Essen, Nr. 165, Febr. 2008.

79 In Abb. 6 findet man entsprechend eine Häufigkeitsverteilung der Kappa-Werte ( $\kappa$ ).

80 Im Jahresbericht 2010 heißt es hierzu: Rund 30% aller Praxen weisen ein Lambda von 1 auf, sodass hier keine Standardisierung mit Lambda erfolgt. 65% aller Praxen weisen ein Lambda von größer als 1, aber kleiner als 1,4 auf. Bei 5% der Praxen liegen höhere Werte von Lambda vor, die zu deutlich geringeren standardisierten Einnahmen führen.

$$(4b) \alpha_i = \frac{\lambda_i - 1}{p_i} + 1.$$

Wenn hierin der Zähler  $\lambda_i - 1$  im Schnitt ungefähr Null ist, dann kann der zugrundeliegende Wert von  $\alpha_i$  (egal, wie groß  $p_i$  ist) nicht wesentlich verschieden von 1 sein.

**Tabelle 9**

 Standardisierungsfaktor  $\lambda$  bei den befragten Arztpraxen (ZiPP)

$\lambda$	n	in vH	$\lambda$	n	in vH
< 1,01	995	29,04	1,41 – 1,53	79	2,04
1,01 – 1,13	1.460	42,62	1,54 – 1,66	39	1,14
1,14 – 1,26	586	17,10	1,67 – 1,79	20	0,58
1,27 – 1,40	222	6,48	> 1,79	25	0,73

Quelle: ZI-Praxis-Panel Jahresbericht 2010, Abb. 5, S. 16

Aus dieser Tabelle ist nicht der Aufschlagsfaktor  $\alpha$  zu ersehen. Aus Gl. 4b folgt jedoch das  $\alpha$  auch bei sehr kleinen Werten von  $\lambda_i - 1$  (von etwa 0,01 bis 0,13 bei der Modalklasse [am stärksten besetzten Größenklasse]) relativ groß sein kann, wenn nur der Privatpatientenanteil  $p_i$  entsprechend klein ist. Aber ohne Kenntnis von  $p_i$  kann von  $\lambda_i$  nicht auf  $\alpha_i$  geschlossen werden. Um eine Vorstellung von den bei der Standardisierung implizit gegebenen  $\alpha$ -Werten zu bekommen war deshalb eine besondere Auswertung der Daten erforderlich. Darauf beruhen Tab. 9 und Abb. 2.<sup>81</sup>

**Tabelle 10**

 Faktor  $\alpha$  bei Arztpraxen ausgewählter Fachrichtungen (ZiPP)  
 (ohne reine GKV-Praxen und Praxen bei denen  $\alpha$  auf 1 gesetzt wurde)\* - Häufigkeiten in vH

	1 - 1,5	1,5 - 2	2 - 2,5	2,5 - 3	3 - 3,5	3,5 - 4	4 - 4,5	> 4,5
Allgemeinmedizin	36,90	31,55	11,31	8,93	4,17	5,36	1,19	0,60
Gynäkologie	3,78	10,08	39,08	13,03	7,14	8,40	11,34	7,14
HNO-Heilkunde	8,0	9,71	29,71	21,71	8,57	10,86	6,86	4,57
Orthopädie	10,97	10,97	34,84	11,61	11,61	6,45	8,39	5,16
Urologie	4,55	9,09	36,36	11,82	10,91	14,55	9,09	3,64

\* Der Anteil dieser Praxen schwankt zwischen 34,4% (bei Allgemeinmedizinern) und 26,7% (bei Gynäkologen) einerseits und 15,5% bzw. 17,3 % bei HNO-Ärzten bzw. Urologen.

Quelle: Daten des ZiPP, Mitteilung von M. Meschenmoser

Diese Tabelle zeigt, dass  $\alpha$  in der Regel nicht sehr groß ist und nicht in der allgemein zitierten Größenordnung eines „Aufschlagsfaktors“, mit dem es allerdings auch nicht vergleichbar ist.<sup>82</sup> Das gilt insbesondere für Allgemeinmediziner (Hausärzte), bei denen auch der Anteil reiner GKV-Praxen (die ohnehin nicht in den Genuss eines  $\alpha > 1$  kommen können) mit 6,25% relativ hoch ist.<sup>83</sup> Auch bei den Gynäkologen ist der Anteil der reinen GKV-Praxen mit 6,46% recht hoch, aber  $\alpha$  scheint doch hier (wie auch bei den anderen in Tab. 9 aufgeführten Fachrichtungen) im Schnitt höher zu liegen als bei den Allgemeinmedizinern. Abb. 2 stellt die Zahlen der

81 Eine ergänzenden Tabelle und Graphik befindet sich im Anhang zu diesem Papier (Tab. A4 und Abb. 3). In den folgenden Fußnoten (insbes. 86f) habe ich einige Detailinformationen verarbeitet, die ich von Herrn Mirko Meschenmoser bekommen habe.

82 Bekannt ist v.a. die Schätzung von  $\alpha$  mit etwa 2,3 nach Wasem et al.

83 Zusammen mit Praxen, bei denen aus bestimmten Gründen ein  $\alpha$  nicht ermittelt werden konnte bzw. auf 1 gesetzt werden musste fallen damit 34,4% aus der Betrachtung aus und von den verbliebenen Praxen haben 36,9% ein  $\alpha$  zwischen 1 und 1,5.

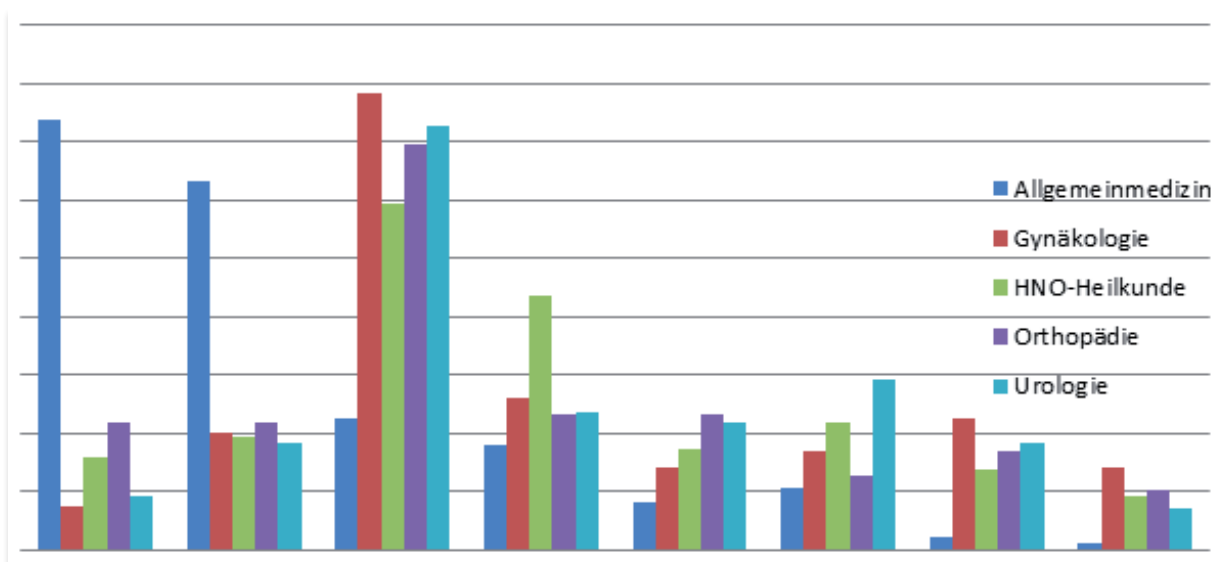
Tab. 10 noch einmal graphisch dar.

Es gibt eine Reihe von Gründen dafür, dass das Ergebnis des ZiPP hinsichtlich  $\alpha$  nicht dahingehend missverstanden werden sollte, dass sich das Vergütungsniveau der PKV kaum von dem der GKV unterscheidet (also  $\alpha$  meist um 1 liegt) oder gar unter dem der GKV liegen kann (wenn  $\alpha < 1$  ist).

## Abbildung 2

Häufigkeitsverteilung der  $\alpha$ -Werte

(ohne reine GKV-Praxen und Praxen bei denen  $\alpha$  auf 1 gesetzt wurde)



Quelle: ZI-Praxis-Panel 2010, Werte siehe Tabelle 10

Nach der Definition von  $\alpha$  als  $e_P/e_G$  sind pro-Kopf Einnahmen bei zwei Arten von Patienten ins Verhältnis zu setzen, aber

- Die Relation  $e_P/e_G$  entspricht nicht dem Aufschlagsfaktor, bei dem PKV- und GKV-Vergütungen für die gleiche Behandlung ins Verhältnis gesetzt werden. Wenn die beiden Patientenarten ganz unterschiedliche Behandlungen (hinsichtlich Zeitaufwand und Kosten) in Anspruch nehmen und den Arzt unterschiedlich oft aufsuchen wird  $e_P/e_G$  (also  $\alpha$  im ZiPP) nicht dem so definierten Aufschlagsfaktor entsprechen. Aber
- aufgrund der erfragten Merkmale (und damit der im ZiPP verfügbaren Daten) und der Nichtvergleichbarkeit der Abrechnungssysteme (PKV vs. GKV) ist auch in vielen Fällen eine korrekte Relation  $e_P/e_G$  nicht zu bilden, sondern nur eine mehr oder weniger gute Näherung an  $e_P/e_G$ .

Die begrenzte Aussagefähigkeit von  $\alpha$  im Rahmen der Definition von  $\lambda$  mindert jedoch nicht den Wert der  $\lambda$ -Methode. Es sollte deshalb noch einmal betont werden, dass die Standardisierung mit der  $\lambda$ -Methode allein auf rein definitorischen Zusammenhängen zwischen Einnahmen (pro Patient), dem Anteil  $p$  der Privatpatienten und der damit gebildeten Relation  $\alpha = e_P/e_G$  aufbaut und keine Modellannahmen über das Verhalten von Ärzten oder Patienten erforderlich macht. Eine davon zu unterscheidende ganz andere Frage ist, ob dieses  $\alpha$  allein Ausdruck der reinen Vergütungsunterschiede ist und damit als Aufschlagsfaktor interpretiert werden kann (und nicht auch Häufigkeit und Art der Inanspruchnahme ärztlicher Leistungen)<sup>84</sup> und ob es die

<sup>84</sup> Wenn mit der Standardisierung sozusagen eine reine GKV-Praxis simuliert werden soll, so ist nicht nur daran zu denken, dass ein Arzt bei Behandlung eines Privatpatienten für die gleiche Art der Behandlung eine höhere Vergütung erhält (was zu messen mit dem „Aufschlagsfaktor“ i.d.R. beabsichtigt wird), sondern auch daran, dass sich Privatpatienten evtl. hinsichtlich der Nachfrage nach ärztlichen Leistungen anders verhalten als Kassenpatienten. Es könnte z.B. sein, dass sich die Altersstruktur und damit auch der Gesundheitszustand der beiden Patientengruppen signifikant unterscheiden, was empirisch zu untersuchen wäre. Jedenfalls sind dies Einflüsse auf die sich im ZiPP ergebenden Werte für  $\lambda_i$  (und damit auch für  $\alpha_i$ ) neben dem sonst meist

Daten überhaupt erlauben, in allen Fällen eine korrekte Relation  $e_P/e_G$  zu bilden (beide Fragen sind wohl zu verneinen).

Man kann gute Gründe nennen, die dafür sprechen, dass man Praxen, die reine GKV-Praxen sind, bzw. bei denen  $\alpha$  auf 1 gesetzt wurde nicht - wie hier geschehen - aus der Betrachtung ausklammern sollte, und dass so der Eindruck eines zu hohen Werts für  $\alpha$  entsteht. Wir präsentieren deshalb im Anhang eine Tabelle und Abbildung analog zu Tab. 10 und Abb. 2. Denn: In welchen Fällen wird  $\alpha = 1$  gesetzt?

Das betrifft vor allem Psychotherapeuten und Psychosomatiker weil entsprechende Leistungen von einigen Privatkassen nur zum Teil bzw. gar nicht erstattet werden und die Differenz zu GOÄ-Vergütung oft nur marginal ist.<sup>85</sup> Es empfiehlt sich dann  $\alpha = 1$  zu setzen um solche Fälle nicht auszuschließen zu müssen.

Zum zweiten oben genannten Punkt (Möglichkeit, die Relation  $e_P/e_G$  mit ZiPP Daten zu bilden) ist zu bemerken: Es ist nicht immer möglich aus den ZiPP Daten  $\alpha$  im Sinne einer der Relation  $e_P/e_G$  von (Durchschnitts-) Einnahmen je Patient korrekt zu bestimmen:

1. Zu Privateinnahmen zählen alle Nicht-GKV-Einnahmen, was nicht gleich zusetzen ist mit Einnahmen aus Abrechnungen mit privaten Krankenversicherungen (PKVen). Diese Einnahmen enthalten neben (nicht mit PKV Vergütungen vergleichbaren) Vergütungen bei Behandlung von Patienten von Berufsgenossenschaften (BG)<sup>86</sup> auch Einnahmen aus IGeL-Leistungen, Vergütungen für Gutachten und ähnliche Positionen.
2. Im ZiPP wird  $\alpha$  praxisspezifisch (also  $\alpha_i$ ) aus GKV-Fallwert und PKV-Fallwert berechnet. Ermitteln lässt sich allerdings nur ein GKV-Fallwert. In dem privatärztlichen Bereich (ambulant) gibt es keine „Fälle“. Bei jeder Inanspruchnahme wird eine Rechnung gestellt und es gibt deshalb auch keinen wirklichen PKV-Fallwert. Im ZiPP wurden auch keine Privat-/BG-Fälle abgefragt sondern nur Angaben zu GKV- und BG-Patientenanteilen. Aus diesen Angaben lässt sich nur ein Pseudo-Privatpatientenanteil ( $p_i$ ) und damit auch nur ein Pseudo-PKV-Fallwert ermittelt, der dann in die Berechnung von  $\alpha$  eingeht (quasi als  $e_P$ ).
3. Anders als bei den im „Wasem-Ansatz“ benutzten Daten, sind im ZiPP keine Daten enthalten, die es erlauben Unterschiede im Leistungsverhalten bei GKV-Patienten und PKV-Patienten abzubilden. Es sind daher durchaus Konstellationen denkbar, bei denen der GKV-Fallwert höher ist, als der (Pseudo-)„PKV-Fallwert“.

Trotz der eingeschränkten Vergleichbarkeit der ZiPP-Werte für die praxisindividuellen  $\alpha_i$  Werte im Rahmen der  $\lambda$  Methode mag es jedoch von Interesse sein, die  $\alpha$ -Werte für Fachgruppen, Regionen usw. zu aggregieren, d.h. hierfür gewogene Mittel zu berechnen. In einer unveröffentlichten Ausarbeitung für das ZI hat der Verfasser hierzu Überlegungen präsentiert. Es bieten sich danach verschiedene Möglichkeiten an, gewogene Mittel der  $\alpha_i$  zu bilden, was dann jeweils auch unterschiedlich gewichtete mittlere Privatpatientenanteile  $p_i$  oder Privatumsatzanteile  $u_i$  impliziert, so dass aggregierten  $\alpha$ -Werte, z.B. für ärztliche Fachgruppen unterschiedlich sinnvoll interpretierbar sind. Ein ungewogenes Mittel  $\bar{\alpha}$  auf Basis der Häufigkeitsverteilungen von Tab. 10 bzw. A4 ist nicht sehr sinnvoll, weil es unterschiedlich gewogene Mittelwerte der  $p$

und  $u$  Werte impliziert. Obgleich für das praxisindividuelle  $\alpha_i$  gilt  $\alpha_i = \frac{u_i}{p_i} \cdot \frac{1-p_i}{1-u_i}$ , wäre es z.B.

nicht sinnvoll das aggregierte  $\alpha$  als  $\bar{\alpha}^* = \frac{\bar{u}}{\bar{p}} \cdot \frac{1-\bar{p}}{1-\bar{u}}$  zu schätzen mit ungewogenen oder (mit  $N_i/$

allein im Vordergrund stehenden Vergütungsunterschieden für ansonsten gleiche Leistungen des Arztes.

85 Beispiel: EBM-Ziffer 35224 Verhaltenstherapie, große Gruppe 20,33 €; GOÄ Ziffer 871 (mit 2,3-fachem Hebesatz 20,10 €).

86 Die Ärzte erhalten für Berufsgenossenschaftspatienten eine Entlohnung nach einer eigenen BG-GOÄ, die sich im Wesentlichen an einer alten Fassung der GOÄ orientiert. Anders als bei Privatpatienten kann hier in der Regel aber nur ein sehr geringer Hebesatz verwendet werden. Die Differenz zu der äquivalenten EBM-Vergütung ist auch meist nur marginal. Andererseits bestehen durchaus Unterschiede zu GKV-Versicherten, da das Leistungsspektrum anders aussieht, und insbesondere häufig Arbeitsunfälle die Ursache für die Inanspruchnahme von ärztlichen Leistungen sind.

$\Sigma N_i$ ) gewogene Mittelwerten der  $p_i$  und

$u_i$  Werten. Es kann hier nicht auf Einzelheiten der möglichen Formeln für ein aggregiertes  $\alpha$  eingegangen werden. Untersucht wurden Möglichkeiten der Gewichtung der  $\alpha_i$  Werte u.a. mit  $p_i/\Sigma p_i$ ,  $N_{pi}/\Sigma N_{pi}$  oder  $N_i/\Sigma N_i$  (so dass Praxen mit relativ mehr Privatpatienten oder auch größere Praxen stärker „ins Gewicht fallen“, d.h. den Mittelwert bestimmen) aber auch  $e_{Gi}/\Sigma e_{Gi}$  oder  $E_i/\Sigma E_i$ . Es zeigte sich, dass es schwierig ist, hier Kriterien für die Wahl der „richtigen“ Aggregation (Mittelwertbildung) anzugeben.

## 5. Zusammenfassung

Das ZiPP soll laufend Informationen liefern über Versorgungsleistungen, Kosten und Erträge von Praxen niedergelassener Ärzte gegliedert nach Regionen und ärztlichen Fachgruppen. Es ist eine Wiederholungsbefragung (Panelerhebung) zur wirtschaftlichen Lage von „Kassenpraxen“ in einem weiten Sinne, die neben detaillierten Angaben zur finanziellen Situation (Kosten, Erträge, Investitionen und deren Finanzierung usw.) auch viele versorgungsrelevante Merkmale umfasst, wie Art und (zeitlicher) Umfang ärztlicher Leistungen, Praxisausstattung (Personal, Räume, Geräte), Organisationsstruktur usw. Die damit gewonnenen Daten sind unmittelbar relevant für Vergütungsentscheidungen über die Aktualisierung des einheitlichen Bewertungsmaßstab (EBM) sowie für Fragen der Sicherstellung der ambulanten ärztlichen Versorgung. Fragen zu Investitionen und subjektiven Einschätzungen gestatten auch Betrachtungen, die über eine Beschreibung des Status quo hinausgehen.

Die amtliche KSE bezweckt demgegenüber mehr eine Erlös-Kosten-Betrachtung eines Wirtschaftszweiges im Rahmen der Volkswirtschaftlichen Gesamtrechnungen (VGR) und zur Information der Allgemeinheit. Sie dient nicht spezifisch der Unterstützung von Honorarverhandlungen der Vertragsärzte und ist weniger detailliert hinsichtlich versorgungsrelevanter Merkmale der Praxen. Andererseits umfasst die KSE auch ausschließlich privatärztliche Praxen, sowie Zahnarzt- und Tierarztpraxen.

Beide Erhebungen sind geschichtete Stichproben, die KSE ist jedoch nicht, wie das ZiPP als laufende, jährlich durchgeführte Panelerhebung konzipiert. In mehrjährigen Turnus erhebt die KSE eine Stichprobe aus den jeweils im Statistischen Unternehmensregister (URS) enthaltenen Praxen (Unternehmen). Die KSE liefert im Unterschied zu einer Panelerhebung keine Daten, die eine explizite Quantifizierung (und Eliminierung) des Einflusses von Strukturveränderungen bzw. -unterschieden auf die ermittelten Kenngrößen (z.B. absolute Größen wie Umsätze oder Verhältniszahlen, wie z.B. die Rentabilität) erlauben.

Datenbasis für die Pflege des URS in der amtlichen Statistik sind primär Angaben der Arbeitsverwaltung sowie andere amtliche statistische Erhebungen. Bei dem im ZiPP als Auswahlgrundlage genutzten Bundesarztregister (BAR) der KBV ist in höherem Maße Vollständigkeit gewährleistet, weil Zulassung und Eintragung in das Register rechtlich gekoppelt sind. Das ZI kann darüber hinaus aber auch noch andere Datenquellen zur Erhebung (bzw. via Verknüpfung mit ZiPP-Daten) und Validierung der ZiPP Erhebungsergebnisse nutzen, die der KSE nicht zur Verfügung stehen. Dies gilt insbesondere für die sog. VDX-Daten (Abrechnungsdaten der KBV).

Eine Besonderheit des ZiPP ist auch die Einbeziehung des Steuerberaters bei einem Teil der Befragung und die Bestätigung aller finanziellen Angaben durch ein Testat des Steuerberaters.

Methodisch besonders bemerkenswerte Unterschiede zwischen ZiPP und KSE sind vor allem

1. das Panelkonzept, d.h. dass im ZiPP eine echte Längsschnittsbetrachtung (Darstellung individueller Abläufe, d.h. Zusammenfügen der Daten der gleichen Praxis im Zeitablauf) zur Darstellung von Entwicklungen frei von Strukturveränderungen
2. wegen der Nutzung des ZiPP zur Anpassung von Arzthonoraren wird im ZiPP eine „Stan-

standardisierung“ der Angaben zu den Erlösen (Umsätzen) und Reinerträgen vorgenommen, die bei der KSE nicht erfolgt; neben den tatsächlichen Angaben werden auch in einem gewissen Sinne strukturbereinigte Werte ausgewiesen.

Bei der Standardisierung geht es um die Bestimmung von Einnahmen einer fiktiven Praxis (oder „Normpraxis“), d.h. um die Einnahmen, die eine Praxis gleicher Größe (gleiche Patientenzahl) gehabt hätte, wenn sie eine reine GKV-Praxis (Privatpatientenanteil in Höhe von  $p = 0$ ) gewesen wäre und eine bestimmte jährlichen Arbeitszeit je Arzt in Höhe von  $A^*$  nicht über- oder unterschritten hätte. Wenn die durchschnittlichen Einnahmen pro PKV-Patient um den Faktor  $\alpha$  größer sind als die durchschnittlichen Einnahmen pro GKV-Patient und die geleistete Arbeitszeit  $A$  größer ist als  $A^*$  bewirkt die Standardisierung dass die fiktiven (standardisierten) Einnahmen geringer sind als die tatsächlichen wenn der Privatpatientenanteil  $p > 0$  ist und die Arbeitszeit  $A > A^*$  ist. Die Standardisierung ist erforderlich, weil zur Orientierung bei Honoraranpassungen im Bereich der Kassenärzte nicht die Einnahmen der tatsächlich beobachteten Arztpraxen, sondern die daraus errechneten Einnahmen einer Praxis mit  $p = 0$  und  $A = A^*$  herangezogen werden.

## Abschließende Übersicht

Hauptkennzeichen der beiden Erhebungen

	KSE (StBA)	ZiPP (ZI)
Zweck	VGR, Information der Öffentlichkeit (alle Nutzer der amtl. Statistik)	Honorarverhandlungen, Sicherstellung der vertragsärztlichen Versorgung
Art der Erhebung	Amtliche Statistik mit Auskunftspflicht, wiederholte Querschnitte (geschichtete Stichprobe aus dem URS); keine Verknüpfung mit externen Daten	Private Statistik, Teilnahme freiwillig, Panelerhebung (geschichtete Stichprobe aus dem BAR); Verknüpfung mit VDX (Abrechnungs-) Daten möglich
Standardisierung	Keine der Standardisierung <sup>a)</sup> entsprechende Umrechnung von Daten	Ausweis tatsächlicher und standardisierter Einnahmen
finanzielle Daten	weniger detailliert hinsichtlich Kosten, keine Investitionen und Angaben zu Vermögen und Finanzierung	detaillierte Angaben zu finanziellen Variablen, Mitwirkung des Steuerberaters (Testat!) bei entspr. Fragen
Praxisdaten	weniger detailliert zur Größe, Versorgungsleistung <sup>b)</sup> etc. der Praxis	mehr Details zur Unternehmensdemographie und zum Leistungsangebot <sup>b)</sup>
Fehler	weniger detaillierte Angaben zum Stichprobenfehler	mehr Nichtbeantwortung (Freiwilligkeit), Angaben zum Stichprobenfehler

a) Standardisierung ist eine „Bereinigung“ (Vergleichbarmachung) unterschiedlicher einnahmerelevanter Strukturen (bzgl. Arbeitszeit und Privatpatientenanteil).

b) Beispielsweise Angaben zu den von den Ärzten vertretenen Fachrichtungen, Anzahl der Behandlungsfälle, Arbeitszeit der Ärzte und des nichtärztlichen Personals, Praxisöffnungszeiten, Ausstattung der Praxis mit Räumen, Geräten etc.

Schließlich ist auch zu bedenken, dass infolge der Freiwilligkeit der Erhebung im Falle des ZiPP mit einer größeren Nichtbeantwortungsquote zu rechnen war (was bei der Stichprobenziehung berücksichtigt wurde). In dem Maße, in dem derartige Antwortausfälle zufällig sind können sie durch „Hochrechnung“ ausgeglichen werden. Anders ist es bei systematischen Ausfällen, insbesondere dann, wenn über unterschiedlichen Verteilungen (Strukturen) der Gruppe der Antwortenden und der Nichtantwortenden wenig bekannt ist.

## Tabellenanhang

### Tabelle A1

Daten zur Stichprobenziehung

Fachgruppe	ZG	KS	KS/ZG	AP	ES	ES/AP
Hausärzte (ohne Kinderärzte)	35.999	474	1,32	2.310	302	13,07
Kinderärzte	3.917	408	10,42	2.026	296	14,61
Anästhesisten	1.626	377	23,19	1.443	131	9,08
Augenärzte	3.341	401	12,00	1.946	191	9,82
Chirurgen	2.087	371	17,76	1.673	225	13,45
Frauenärzte	6.883	432	6,28	2.166	439	20,27
HNO-Ärzte	2.641	385	14,59	1.773	278	15,68
Phoniatry und Pädaudiologie	15	14	94,96	16	6	37,50
Hautärzte	2.326	373	16,06	1.668	177	10,61
Internisten - SP Angiologie	71	60	83,92	63	20	31,75
Internisten - SP Endokrinologie	54	44	82,12	55	7	12,73
Internisten - SP Gastroenterologie	314	173	55,12	320	102	31,88
Internisten - SP Hämato-/Onkologie	193	130	67,16	184	29	15,76
Internisten - SP Kardiologie	638	250	39,21	649	94	14,48
Internisten - SP Nephrologie	319	187	58,53	293	7	2,39
Intern. - SP Pneumologie (inkl. Lungenärzte)	541	237	43,83	548	66	12,04
Internisten - SP Rheumatologie	167	114	68,46	168	25	14,88
Internisten – Sonstige	1.060	318	30,03	1.101	148	13,44
Kinder- und Jugendlichenpsychiater	507	219	43,17	510	106	20,78
Mund-Kiefer-Gesichts-Chirurgen	751	0	0	0	0	0
Nervenärzte ohne Neurologen/Psychiater	1.915	356	18,61	1.561	186	11,92
Neurologie	257	155	60,47	263	81	30,80
Psychiatrie	805	262	32,57	829	200	24,13
Neurochirurgie	191	127	66,69	193	5	2,59
Orthopäden	3.018	403	13,35	1.827	248	13,57
Pathologen	280	157	56,19	272	7	2,57
Radiologen	633	259	40,95	620	57	9,19
Strahlentherapie	61	52	84,59	57	5	8,77
Nuklearmediziner	155	112	72,45	146	12	8,22
Urologen	1.655	351	21,20	1.548	182	11,76
Humangenetik	34	29	84,74	37	2	5,41
Physikalische u. rehabilitative Medizin	200	133	66,70	220	32	14,55
Ärztl. u. psych. Psychotherapeuten	15.678	442	2,82	2.215	531	23,97
Psychosomatische Med. u. Psychotherapie	2.063	337	16,33	1.467	283	19,29
Fachübergr. Gemeinschaftspraxen	2.745	384	13,98	1.648	162	9,83
MVZ	1.454	344	23,63	18	19	105,56
<b>Insgesamt</b>	<b>94.594</b>	<b>8.873</b>	<b>9,38</b>	<b>31.833</b>	<b>4.661</b>	<b>14,64</b>

Abkürzungen: **ZG** = Zielgesamtheit (Auswahlrahmen; Praxen im BAR); **KS** = kalkulierte Stichprobe; **AP** = Anzahl der angeschriebenen Praxen; **ES** = effektiv erreichte Stichprobe (Rücklauf); **KS/ZG** = kalkulierte Stichprobe in Prozent der Zielgesamtheit; **ES/AP** entsprechend



**Tabelle A2**

Praxen der Grundgesamtheit differenziert nach den 108 Schichten im ZiPP

Fachgruppe	KS Kernstädte	LR Ländl. Region	VK Verdicht- tete Kreise	Gesamt
Hausärzte (ohne Kinderärzte)	11.029	6.598	18.372	35.999
Kinderärzte	1.415	578	1.924	3.917
Anästhesisten	818	148	660	1.626
Augenärzte	1.289	514	1.538	3.341
Chirurgen	838	323	926	2.087
Frauenärzte	2.738	1.011	3.134	6.883
HNO-Ärzte	1.084	383	1.174	2.641
Phoniatry und Pädaudiologie	10	0	5	15
Hautärzte	978	300	1.048	2.326
Internisten - SP Angiologie	32	6	33	71
Internisten - SP Endokrinologie	43	5	6	54
Internisten - SP Gastroenterologie	158	29	127	314
Internisten - SP Hämato-/Onkologie	93	22	78	193
Internisten - SP Kardiologie	271	71	296	638
Internisten - SP Nephrologie	115	70	134	319
Internisten - SP Pneumologie (inkl. Lungenärzte)	236	73	232	541
Internisten - SP Rheumatologie	88	12	67	167
Internisten - Sonstige	308	243	509	1.060
Kinder- und Jugendlichenpsychiater	261	41	205	507
Mund-Kiefer-Gesichts-Chirurgen	375	52	324	751
Nervenärzte ohne Neurologen/Psychiater	904	229	782	1.915
Neurologie	123	23	111	257
Psychiatrie	479	67	259	805
Neurochirurgie	98	19	74	191
Orthopäden	1.216	418	1.384	3.018
Pathologen	144	16	120	280
Radiologen	249	106	278	633
Strahlentherapie	29	2	30	61
Nuklearmediziner	67	21	67	155
Urologen	642	244	769	1.655
Humangenetik	31	0	3	34
Physikalische und rehabilitative Medizin	92	18	90	200
Ärztliche u. psychologische Psychotherapeuten	8.751	1.097	5.830	15.678
Psychosomatische Medizin und Psychotherapie	1.196	126	741	2.063
Fachübergreifende Gemeinschaftspraxen	1.105	345	1.295	2.745
MVZ	700	254	500	1.454
<b>Insgesamt</b>	<b>38.005</b>	<b>13.464</b>	<b>43.125</b>	<b>94.594</b>

**KS** = Kreistypen 1 und 5; **LR** = Kreistypen 4, 7 und 9; **VK** = Kreistypen 2, 3, 6 und 8

**Tabelle A3**

Differenzierung nach Arztgruppen in der KSE und im ZiPP

Das ZiPP sieht insgesamt 36 verschieden gestaltete Fragebögen (Masterfragebögen) vor, weil die (räumlichen, technischen und personellen) Verhältnisse der Praxen in den verschiedenen Fachgruppen unterschiedlich sind. Bei der KSE wird nach unserer Kenntnis in dieser Hinsicht nur unterschieden nach den oben genannte „Branchen“ (85.12 Arztpraxen, 85.13.0 Zahnarztpraxen usw.). Die Differenzierung nach Fachgruppen erfolgte bei der KSE aufgrund der in freier Wortwahl erfolgten Eigenangabe von Ärzten, die dann verschlüsselt wurde, wobei dann von Amts wegen die Differenzierung gem. der folgenden Tabelle vorgenommen wurde.

KSE (StBA)	ZiPP (ZI)
1. Allgemein-/Praktische Ärzte	Hausärzte
2. Internisten	Innere Medizin, SP Endokrinologie Innere Medizin, SP Gastroenterologie Innere Medizin, SP Angiologie Innere Medizin, SP Kardiologie Innere Medizin, SP Pneumologie Innere Medizin, SP Nephrologie Innere Medizin, SP Hämatologie/Onkologie Innere Medizin, Sonstige
3. Frauenheilkunde	Gynäkologie
4. Kinderheilkunde	Kinder- und Jugendmedizin
5. Augenheilkunde	Augenheilkunde
6. Hals-Nasen-Ohrenheilkunde	Hals-Nasen-Ohrenheilkunde
7. Orthopädie	Orthopädie
8. Chirurgie	Chirurgie Neurochirurgie Mund-Kiefer-Gesichtschirurgie
9. Haut-/Geschlechtskrankheiten	Hautärzte
10. Radiologie und Nuklearmedizin	Radiologie Strahlentherapie Nuklearmediziner
11. Neurologie, Psychiatrie	Nervenärzte ohne Neurologen/Psychiater Ärztl. und psycholog. Psychotherapeuten Psychosomatische Medizin u. Psychotherapie Kinder- u. Jugendpsychologie/-psychotherapie Psychiatrie Neurologie
12. Urologie	Urologie
unklare Zuordnung	Anästhesiologie Pathologie Phoniatrie und Pädaudiologie Physikalische u. Rehabilitative Medizin Humangenetik Fachübergreifende Gemeinschaftspraxen

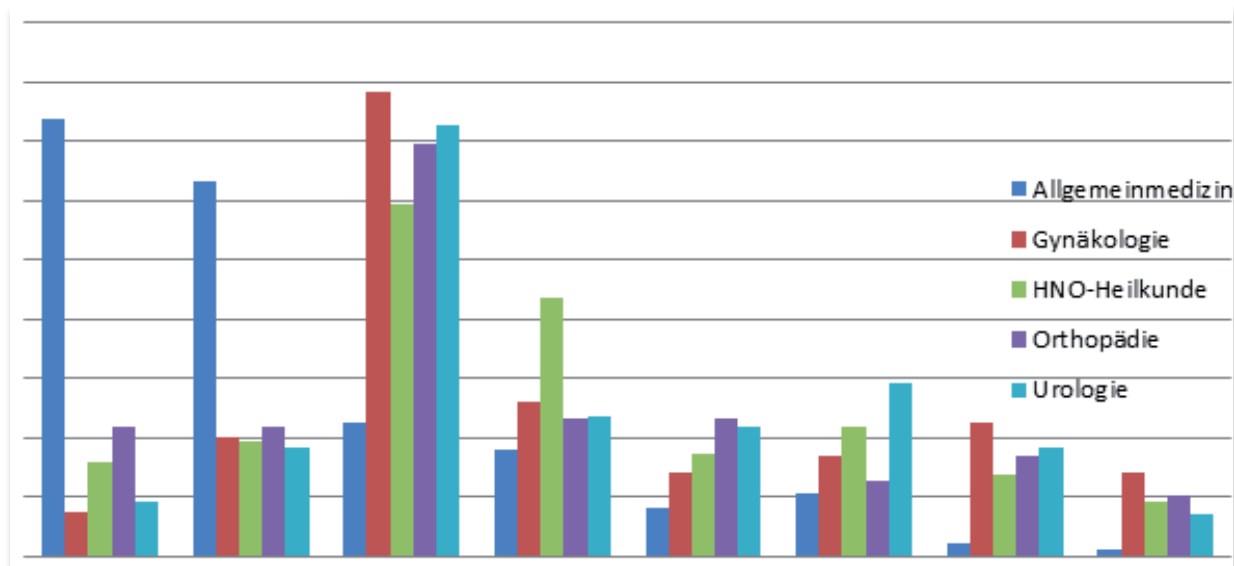
**Tabelle A4**

Faktor  $\alpha$  bei Arztpraxen ausgewählter Fachrichtungen  
 (alle Praxen umfassend, auch in Spalte 1 reine GKV-Praxen und Praxen bei denen  $\alpha = 1$  gesetzt wurde)

	1	1-1,5	1,5 - 2	2 - 2,5	2,5 - 3	3 - 3,5	3,5 - 4	4 - 4,5	> 4,5
Allgemeinmedizin	34,38	24,29	20,70	7,42	5,86	2,73	3,52	0,78	0,39
Gynäkologie	26,77	2,77	7,38	28,62	9,54	5,23	6,15	8,31	5,23
HNO-Heilkunde	15,46	6,76	8,21	25,12	18,36	7,25	9,18	5,80	3,87
Orthopädie	20,51	8,72	8,72	27,69	9,23	9,23	5,13	6,67	4,10
Urologie	17,29	3,76	7,52	30,08	9,77	9,02	12,03	7,52	3,01

**Abbildung 3**

Häufigkeitsverteilung der  $\alpha$ -Werte  
 (mit reinen GKV-Praxen und Praxen bei denen  $\alpha$  auf 1 gesetzt wurde)



Quelle: ZI-Praxis-Panel 2010, Werte siehe Tabelle A4