



Andreas Wähnert

[Zentrales Prüfungsdienste-Finanzamt (ZPD) Kiel]

1. öffentliche Auflage (11.1.2019)

10. überarbeitete interne Auflage (Erstauflage v. 29.11.2007)

 andreas.waehner@fa-zpd.landsh.de

Bei dem Handbuch für die Summarische Risikoprüfung (SRP) handelt es sich nicht um einen Erlass des Finanzministeriums Schleswig Holstein sondern um eine Hilfestellung für Prüferinnen und Prüfer für den Einsatz moderner Prüfungsmethoden. Nicht alle Ausführungen des Handbuchs stellen daher zwingend die Position des Finanzministeriums Schleswig-Holstein dar.

Inhalt:	Seite
1. SRP-Kurzbeschreibung	6
2. Intention der SRP (Kernaspekte)	11
2.1. Heutige Vollzugssituation und Steuerausfallrisiken	14
2.2. Moderne Prüfungsmethodik in der digitalisierten Welt	18
2.3. Vorteile mehrperspektivischer Prüfungsnetze wie SRP	22
3. Einbindung der SRP in Prüfung und Risikomanagement	26
3.1. Die SRP als Orientierungshilfe zu Prüfungsbeginn	28
3.2. Die SRP im besonderen Kontext des § 158 AO	30
4. Die Verwendung der SRP-Vorlagen (Basis-SRP)	33
4.1. Formalprüfung: Ordnungsmäßige (prüfbare) Aufzeichnungen	38
4.2. Regelbasierte Ansätze (Zeitreihenanalysen)	47
4.2.1. Betriebsprofil	52
4.2.2. Prüfungssystematik der SRP-Vorlage BWA	57
4.2.3. Jahressummen	62
4.2.4. Quartalsrelationen	63
4.2.5. Monatszeitreihen	66
4.2.5.1. Doppelskaliert größenangepasster Zeitreihenvergleich	68
4.2.5.2. Einzelskalierter Zeitreihenvergleich	70
4.2.5.3. Verhältniszeitreihe (z.B. Aufschlagsatzzeitreihen)	72

4.2.6. Gleitschlittenzitreihen	73
4.2.7. Entwicklungsvergleich der einzelnen Datenreihen	76
4.2.8. Rückrechnung der monatlichen Warenbestände	79
4.2.9. Interaktive Optionen zur verbesserten Sachaufklärung	82
4.2.9.1. Einzelkorrektur von Datensätzen	83
4.2.9.2. Einschränkung des Betrachtungszeitraums	84
4.2.9.3. Verschieben der ersten Datenreihe	85
4.2.9.4. Beseitigung einseitiger (konjunktureller) Preiseinflüsse	86
4.2.9.5. Niveaueinpassung eines Grafen	88
4.2.10. Kosten- und weitere Zeitreihen	89
4.2.11. Beweiskraft von (regelbasierten) Zeitreihenanalysen	93
4.3. Musterbasierte Analysen (Häufigkeitsauswertungen)	97
4.3.1. Ablauf von Häufigkeitsanalysen (Chi-Quadrat-Test)	103
4.3.1.1. Beobachtungen i.R. von Häufigkeitsanalysen	103
4.3.1.2. (Punkt-)Erwartungen i.R. von Häufigkeitsanalysen	105
4.3.1.3. Übereinstimmungsprüfung i.R. von Häufigkeitsanalysen	107
4.3.1.4. Schlussfolgerungen i.R. von Häufigkeitsanalysen	110
4.3.2. Ziffernanalyse	111
4.3.2.1. Anwendungshinweise zu SRP-Ziffernanalysen	112
4.3.2.2. Einerstellenuntersuchung	115
4.3.2.2.1. Die (Punkt-)Erwartung „Gleichverteilung“	115
4.3.2.2.2. Nicht-manipulative Ursachen für Einerstellenauffälligkeiten	116
4.3.2.2.3. Schrittweise Abklärung von Einerstellenauffälligkeiten	117

4.3.2.3. Zweitziffernanalyse (Benford-Gesetz)	120
4.3.2.3.1. Die (Punkt-)Erwartung „Benford-Gesetz“	120
4.3.2.3.2. Nicht-manipulative Ursachen für Zweitziffernauffälligkeiten	122
4.3.2.3.3. Schrittweise Abklärung von Zweitziffernauffälligkeiten	123
4.3.3. Klassifizierung (Schichtung) der Werte	126
4.3.3.1. Anwendungshinweise zu SRP-Werteklassifizierungen	128
4.3.3.2. Abklärung von Zweitziffernauffälligkeiten mit 99-er Schichtung	132
4.3.3.3. Vorbereitende Werteklassifizierungen für die Verteilungsanalyse	134
4.3.4. Verteilungsanalyse	137
4.3.4.1. Die (Punkt-)Erwartung „Logarithmische Normalverteilung“	139
4.3.4.2. Anwendungshinweise zur SRP-Verteilungsanalyse	142
4.3.4.3. Ergebnisverdeutlichung bei auffälligen Verteilungsanalysen	147
4.3.4.4. Nicht-manipulative Ursachen für auffällige Verteilungsanalysen	150
4.3.4.5. Abklärung von auffälligen Verteilungsanalysen	151
4.3.4.6. Quantifizierung von auffälligen Verteilungsanalysen	153
4.3.5. Beweiswert von (musterbasierten) Häufigkeitsanalysen	155
4.4. Gesamtergebnis einer plausibilitätsorientierten SRP	160
4.5. SRP im Verhältnis zu anderen Prüfungsmethoden	163
4.6. Welches Beweismaß fordert § 158 AO?	165
4.7. Einleitung des Strafverfahrens nach § 10 BpO	170
4.8. Schätzung im Besteuerungsverfahren (Quantilsschätzung)	173
4.9. Strafrechtliche Schätzung	190

5. Freie SRP-Anwendung (Systematische Visualisierung)	194
5.1. Einsatzbereiche systematischer Visualisierung	195
5.2. Visualisierende Prüffeldsuche	198
5.3. Visualisierte Übereinstimmung (Schnittstellenprüfung)	205
5.4. Visualisierte Schlüssigkeit („sprechende Diagramme“)	209
5.5. Ausblick zur visualisierenden Datenanalyse	214
6. (Ausgewählte) Literatur zur „Neuen Prüfungstechnik“	217

1. SRP-Kurzbeschreibung

Die **Summarische Risikoprüfung (SRP)** ist

- eine vorlagenunterstützte (technisch entlastende und bedienerführende) **Filterprüfung** für elektronische Betriebs-/Besteuerungsdaten (Buchführung, Vorsysteme etc.),
- um mit Hilfe mehrperspektivisch vernetzter Gesamtbildverfahren (**Prüfungsnetz**)
- allgemein einzelfallbezogene **Risiko-Prüffelder** systematisch festzustellen und einzugrenzen
- sowie speziell beim Prüffeld „Einnahmenvollständigkeit“ eine möglichst umfassende und abgesicherte **Schlüssigkeitsbeurteilung** – i.S. einer möglichst hohen Aussagegewissheit – vornehmen zu können;
- dadurch soll ein **ressourcenschonendes Risikomanagement** mit zeitgemäßen (digitalen) Prüfungsmethoden unterstützt werden.

In Anbetracht der großen Anzahl an Unternehmen in Deutschland reichen die vorhandenen Personalkapazitäten nur für **unbefriedigende Prüfungsquoten**¹ im Bereich der Kleinst- bis Mittelbetriebe, so dass infolge der (statistischen) Kurzlebigkeit viele von diesen nie intensiv geprüft werden. Deshalb könnte die Außenprüfung z.B. bei einem Kleinbetrieb nach Literaturmeinung² an sich schon als Akt staatlicher Willkür empfunden werden. Die Veranlagungsprüfung kann eine vergleichbare Tiefe nicht erzielen, weil der digitale Zugriff auf die Einzeldaten nach § 147 Abs. 6 AO auf die Außenprüfung beschränkt ist.

Zusätzlich wird die dadurch begrenzte Aufdeckungswahrscheinlichkeit von Steuerdelikten infolge der Verbreitung immer **komfortablerer Manipulationswerkzeuge** wie z.B. Zappern für Warenwirtschaftsprogramme von Apotheken³ weiter reduziert. Dementsprechend wird beim BFH „auf Seiten der Finanzverwaltung ein durchaus nachvollziehbares Bedürfnis, moderne Prüfungsmethoden [...] einzusetzen“⁴ gesehen. Der Bundesrechnungshof

¹ Monatsbericht des BMF für den November 2017 mit der Betriebsprüfungsstatistik 2016, die Prüfungsquoten von 1,0 % für Kleinst-, 3,2 % für Klein- und 6,4 % für Mittelbetriebe ausweist; nachzulesen unter: <https://www.bundesfinanzministerium.de/Monatsberichte/2017/11/Inhalte/Kapitel-3-Analysen/3-7-Ergebnisse-der-steuerlichen-Betriebspruefung-2016.html>.

² Tipke, „Die Steuerrechtsordnung“, Verlag Dr. Otto Schmidt, Bd. III, 1407.

³ Nachzulesen unter: www.apotheke-adhoc.de/nachrichten/nachricht-detail/lauer-fischer-zapper-ist-nicht-von-uns.

⁴ Urt. v. 25.3.2015, BStBl. II 2015, 743, Rz. 54.

erkennt „durch immer komplexere Manipulationssoftware [...] bei der Besteuerung von Bargeschäften ein strukturelles Vollzugsdefizit“⁵. Die Kleine Anfrage „Schaden durch Betrug mit manipulierten Kassensystemen und mögliche Einführung der INSIKA-Lösung zur Betrugssicherung“⁶ zeigt, dass auch die Politik alarmiert ist. Entsprechend erreicht das Thema in den Populärmedien mit Titeln wie „Steuerhinterziehung: Zu Gast bei Betrügern“⁷ eine starke Präsenz und wird an prominenter Stelle vor einem gesellschaftsbedrohenden Verfall der Steuermoral gewarnt⁸.

Unabhängig von vorsätzlichen Eingriffen in Besteuerungsdaten offenbaren gezielte Prüfungsmaßnahmen z.T. gravierende **Datenmängel der elektronischen Betriebsverwaltungssysteme**⁹, die heutzutage hochgradig automatisiert ablaufen – tlw. mit unternehmensspezifischen Anpassungen und intransparent unvollständigen Bedienerrollen. Demnach scheinen die betriebsinternen Kontrollsysteme (IKS) die Risiken von Fehlschlüssen, Datenschnittstellen etc. derzeit nicht ausreichend abzusichern. Welche verheerenden Auswirkungen das Fehlen effektiver Kontrollen haben kann, ist durch die Finanzkrise offengelegt worden.

Besonders im Bereich der **Großbetriebe und Konzerne** kommen Anzeichen weiterer Risiken für dauerhafte Steuerausfälle hinzu. So moniert das Deutsche Institut für Wirtschaftsforschung (DIW) illegale Steuervermeidungsstrategien internationaler Unternehmensverbünde¹⁰, wogegen die groß angelegte OECD-Aktion „Base Erosion and Profit Shifting (BEPS)“¹¹ ansetzen soll. Ohne wirksame Überprüfungsmöglichkeiten der tatsächlichen Verhältnisse können der internationale Informationsaustausch und Regulierungen der Dokumentationsweise allerdings keine ausreichende Sicherung der Steueransprüche gewährleisten.

Zusammenfassend hat die von der Wirtschaft vorangetriebene Digitalisierung und Globalisierung unserer Gesellschaft bekannte Steuerrisiken verschärft und neue geschaffen. Deshalb

⁵ n-tv online v. 25.6.2015, nachzulesen unter: <http://www.n-tv.de/wirtschaft/Laender-kaempfen-gegen-Schummel-Kassen-article15373191.html>.

⁶ Nachzulesen unter: <http://dip21.bundestag.de/dip21/btd/18/061/1806118.pdf>.

⁷ Zeit online v. 20.2.2014, nachzulesen unter: <http://www.zeit.de/2014/09/steuerhinterziehung-gastronomie>.

⁸ Köhler bei der Verleihung des Max-Weber-Preises für Wirtschaftsethik 2008: „Wenn eine Mehrheit es für clever hält, das Finanzamt zu betrügen, dann gehen nicht nur Steuereinnahmen verloren, dann gerät auch der Rechtsstaat und das Vertrauen in die demokratische Ordnung in Gefahr“, nachzulesen unter: www.stern.de/politik/deutschland/manager-schelte-koehlers-neues-lieblingsthema-621740.html. sowie „Steueroase Deutschland“, Süddeutsche.de v. 7.11.2013 und Spiegel 40/2013, 74.

⁹ *Webel/Danielmeyer*, „Schnittstellenverprobung elektronischer Betriebsverwaltungen – das kommende Standardprüffeld?“, StBp 2015, 353.

¹⁰ n-tv online v. 28.5.2013, nachzulesen unter: www.n-tv.de/wirtschaft/Fiskus-entgehen-jedes-Jahr-Milliarden-article10717336.html.

¹¹ Nachzulesen unter: <http://www.oecd.org/ctp/beps.htm>.

gibt es in vielen Finanzverwaltungen Bestrebungen, ein **wirksameres Risikomanagement** aufzubauen.¹² Schwerpunkte der Maßnahmen liegen in den Bereichen „Organisation“ (z.B. Neuaufnahmestellen zur frühzeitigen Risikoerkennung), „Informationsaustausch und Datenbankennutzung“ sowie „automatisierte Filternetzwerke für die Steueranmeldungs- und -erklärungszahlen“. Zum einen sind Quartals- und Jahreswerte für eine wirksame Risikoerkennung jedoch zu sehr verdichtet (und die Risikospuren dadurch „geglättet“), weshalb z.B. Ungarn für das System RADAR differenzierte Rückmeldungen aus durchgeführten Außenprüfungen verkennziffert und in die Risikokalierung einbezieht. Zum anderen verkennt der Einsatz „künstlicher Intelligenz“¹³, dass die Konzentration auf selbstlernende Algorithmen die Prüfungskompetenzen des Personals „Intuition“ und „Erfahrung“ ausblendet und damit wichtige Ressourcen ignoriert.

Wegen des Vollzugriffs auf elektronische Besteuerungsunterlagen nach § 147 Abs. 6 AO liegen in der **Optimierung der Außenprüfung** die erfolgversprechendsten Chancen, insbesondere auch die erhöhten (digitalen) Steuerausfallrisiken einzudämmen. So sieht die Wissenschaft einen Schlüssel für die erfolgreiche Durchführung von Prüfungsmaßnahmen in der Prüfungstechnik mit einem „hohen Maß an Methodenwissen, Flexibilität und Kreativität“¹⁴. Der sinnvoll kombinierte Einsatz der Analysesoftware IDEA[®] mit Tabellenkalkulationsprogrammen wie MS Excel[®] erreicht ein entsprechendes Maß an Funktionalität, um von einem Gesamtüberblick über Besteuerungsdaten zur gezielten Identifikation prüfungsrelevanter Sachverhalte vorzugehen.

Die **SRP** nutzt mit ihrem zentralen Aspekt der „systematischen Visualisierung“ den bekannten wahrnehmungsphysiologischen Vorteil¹⁵ der Verbildlichung von Zahlen und komplexen Sachverhalten in ihren bedienerfreundlich vorkonfektionierten Vorlagen, welche den Einstieg in die digitale Prüfungstechnik erleichtern und dennoch die Entscheidungen bei den PrüferInnen belassen. Ein bundesweites Austauschforum für Prüfungsideen¹⁶ ermöglicht zudem

¹² S. z.B.: Risk Management Guide, Leitfaden – Risikomanagement für Steuerverwaltungen, Europäische Kommission, Generaldirektion Steuern und Zollunion, Indirekte Steuern und Steuerverwaltung, Verwaltungszusammenarbeit und Bekämpfung des Steuerbetrugs, Fiscalis-Projektgruppe Risikoanalyse, Version 1.02, Februar 2006.

¹³ Schwerpunktthema u.a. beim 14. Deggendorfer Forum zur digitalen Datenanalyse mit dem Titel „Die Welt ist komplex: Verstehen Computer sie besser?“, nachzulesen unter: <https://www.dfdda.de/index.php/de/veranstaltungen/forum/14-forum-2018>.

¹⁴ Nachzulesen unter: www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik.

¹⁵ *Barnard* in einer US-amerikanischen Zeitschrift der Werbebranche 1921: „One look is worth a thousand words“, nachzulesen unter: www.oai.de/de/component/content/article/45-publikationen/sprichwort/811-ein-bild-sagt-mehr-als-tausend-worte.html.

¹⁶ *Intranet-Adresse: http://wiki.ktz.testa-de.net/cgi-bin/twiki/view/PRUEFUNGSD/SummarischeRisikopruefung*.

die zeitnahe Verbreitung von Weiterentwicklungen und fördert damit die digitale Prüfungsweise insgesamt.

Der vorlagenunterstützte Bereich der SRP besteht aus zwei Teilen: Erstens identifiziert eine mehrperspektivisch regelbasierte Auswertung von (abhängigen) Betriebsgrößen (z.B. Warenfluss, umsatzbezogene Kosten, Datenschnittstellen oder Vertragsumsetzungen) Prüffelder als „Regelstörungen“ in Zeitreihenanalysen und kann zur Beurteilung der kaufmännischen Schlüssigkeit herangezogen werden. Zweitens suchen systematisch verknüpfte musterbasierte Häufigkeitsuntersuchungen allgemein nach „Musterstörungen“ als prüfungsrelevante Vorgänge (Gestaltungen und Fehler) und speziell in manipulationsanfälligen Betriebsdaten (besonders Tageskasseneinnahmen) nach eindeutigen [stochastisch (mit Hilfe der Wahrscheinlichkeitstheorie) abgesicherten] Eingriffsspuren.

Beide Bereiche sind mit Hilfe der SRP-Vorlagen BWA und Stochastik schnell und variabel handhabbar. Sie setzen nur geringe technische Fähigkeiten und durchschnittliche mathematische Kenntnisse voraus, indem z.B. Anwendungsregeln in der Programmierung bzw. dem Prüfungspfad berücksichtigt werden. In der SRP-Vorlage BWA wird durch die multiple Betrachtung die Gefahr von Interpretationsfehlern (z.B. als Folge nicht beachteter betrieblicher Besonderheiten) reduziert. Die SRP-Vorlage Stochastik konzentriert sich bei der Suche nach Manipulationsspuren zur Vermeidung von Fehldeutungen auf die praxisrelevantesten Hinterziehungsarten und unterstützt das Abklären betrieblicher Ursachen. Beide Vorlagen bieten Möglichkeiten für nachvollziehbare Korrekturen (Interaktivität), um eine zeitgemäße Prüfungstiefe und eine höchstmögliche Absicherung der Ergebnisse zu gewährleisten. Zudem erleichtern sie eine nachvollziehbare Dokumentation der Auswertungen. Zwei integrierte Schätzungsansätze („Quantilsschätzung“ und „Verteilungsanalyse“) sollen zusätzliche Praxisvorteile für die Ermittlung einer Schätzungshöhe bieten – *bei vorliegender Schätzungsnotwendigkeit*.

Die bisherigen **Erfahrungen zur SRP** aus mehreren tausend Prüfungen durch mehr als 2.000 geschulte PrüferInnen in drei Viertel der Bundesländer sowie von der Bundesprüfung, an der Bundesfinanzakademie, beim Zoll, in der niederländischen Finanzverwaltung und in einer Kooperation mit dem Internationalen Währungsfond (IWF) haben gezeigt, dass

- ✓ nach der einwöchigen Basisfortbildung (und zeitnaher Anwendung an eigenen Fällen) üblicherweise ein ausreichendes Methodenverständnis bei den PrüferInnen erreicht wurde,
- ✓ der erforderliche Zeitaufwand für die vorlagenunterstützte SRP als Orientierungshilfe zu Beginn einer Außenprüfung wenige Stunden beträgt,

- ✓ selbst bei nicht vollständiger Durchführbarkeit aller Einzelmethoden (z.B. infolge mäßiger Aufzeichnungsqualität) eine qualifizierte Einschätzung der Vertiefungsnotwendigkeit und Festlegung von Prüffeldern möglich ist,
- ✓ beim bargeldintensiven Einzelhandel in der Mehrzahl der Fälle Implausibilitäten angezeigt werden, was die zeitgemäße Aufdeckungswahrscheinlichkeit erkennen lässt,
- ✓ bei (offensichtlich) manipulierten Unterlagen i.d.R. diverse Auffälligkeiten auftreten, womit die (indizielle) Beweiswirkung gem. BGH-Rechtsprechung erheblich verstärkt wird,
- ✓ konventionelle Ergebniskontrollen z.B. mit Hilfe von (zeitaufwendigeren) Kalkulationen und Durchsuchungsmaßnahmen die Ergebnistendenz regelmäßig bestätigen und
- ✓ die (freie) Anwendung mehrperspektivisch systematischer Visualisierung bei der Suche nach und Abarbeitung von Prüffeldern zu schnelleren, vollständigeren und besser argumentierbaren Resultaten führt.

Von der Prüfungsbedeutung der Verbildlichung wird inzwischen medial berichtet.¹⁷ Sowohl die SRP als auch wichtige Teilaspekte wie das Betriebsprofil oder die systematische Vernetzung sind von der **Kommentierung/Literatur** übernommen worden.¹⁸ Die Aussagekraft (sorgfältig erarbeiteter) Ergebnisse wird z.T. sehr positiv bewertet.¹⁹

Trotz der systematischen Vorlagenunterstützung der SRP mit Maßnahmen zum Fehlerausschluss und zur Ergebnisabsicherung ist ein eingehendes Studium der angewandten Verfahren inkl. kombinierter Beurteilung notwendig und zentrales Anliegen der SRP, damit die Anwendung bis hin zur Schätzung verantwortungsvoll und sachgerecht erfolgt. Dieses **Handbuch**²⁰ hat zwar den Charakter einer Bedienungsanleitung für die SRP-Vorlagen, wobei aber alle vorgestellten Auswertungen mit frei zugänglichen Programmen nachvollzogen werden können. In den ersten beiden Kapiteln werden die Intention der SRP aus Sicht des Autors und eine sinnvolle Einbindung in die einzelne Prüfung sowie das bestehende Risikomanagement beschrieben.

¹⁷ Wirtschaftswoche v. 21.5.2015: „Fiskus 4.0: Mit Big Data auf Steuerfang“, nachzulesen unter: <http://www.wiwo.de/politik/deutschland/fiskus-4-0-mit-big-data-auf-steuerfang/11770850.html>.

¹⁸ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 16 u. 20, bzw. *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24.

¹⁹ Vgl. z.B.: *Pump/Heidl*, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenbuchführung“, StBp 2014, 163: „Der weitere Vorteil von SRP besteht darin, dass die finanzamtliche Beweisführung für die Hinzuschätzung auch im Steuerstrafverfahren gerichtsfest ist.“

²⁰ Die z.T. häufigere Wiederholung von bestimmten Aussagen und insbesondere Fußnoten soll auch die auszugswise Nutzung als Handbuch gewährleisten.

2. Intention der SRP (Kernaspekte)

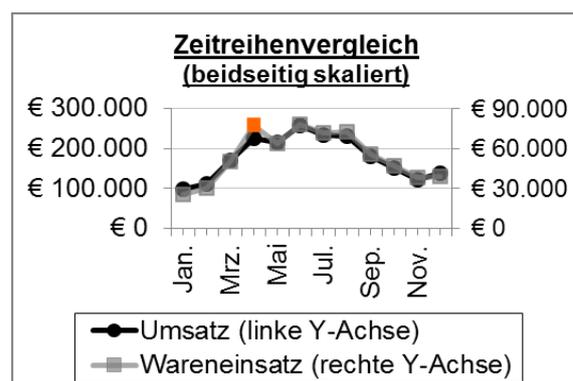
Leitmotive bei der (Weiter-)Entwicklung der SRP als systematisch-mehrperspektivisches Prüfungsnetz sind bzw. waren es,

- die digitalen Vorteile „Rechengeschwindigkeit“, „Variabilität“ und „Interaktivität“ dafür zu nutzen, zu Beginn einer Prüfungsmaßnahme eine **qualifiziertere Risikoeinschätzung** (mit zeitgemäßer Aufdeckungswahrscheinlichkeit betriebsspezifischer Prüffelder) sowie eine möglichst hochgradige Ergebnisklarheit im Rahmen von Plausibilitätsprüfungen zu erreichen, und
- die wahrnehmungsphysiologische Überlegenheit der Verbildlichung beim Erfassen, Bewerten und Kommunizieren komplexer Informationen für eine **schnellere, vollständigere und besser argumentierbare Abarbeitung einzelner Prüffelder** wie z.B. bei der Aufdeckung gesetzeswidriger Gestaltungen (s. [Pkt 5.4.](#)) einzusetzen.

Dazu setzt die SRP zentral auf die **Visualisierung von Steuerungsdaten**, denn

Warenflussdaten		
	Umsatz	Wareneinsatz
Jan.	97.791,65	25.159,28
Feb.	112.294,21	30.485,25
Mrz.	170.895,79	50.337,68
Apr.	226.075,89	77.657,37
Mai	216.656,23	64.092,84
Jun.	257.442,80	77.851,63
Jul.	234.162,84	71.635,76
Aug.	231.841,05	72.596,78
Sep.	181.955,20	55.737,06
Okt.	151.803,76	46.769,50
Nov.	122.264,11	38.390,08
Dez.	137.304,88	38.822,04

„ein Bild sagt mehr als tausend Worte [bzw. Zahlen]“²¹. Das Sprichwort gilt zwar als Erkenntnis der Werbebranche, ist aber wahrnehmungsphysiologisch unumstritten und lässt sich am Eigenversuch einfach nachvollziehen: Während sich aus der Tabelle zu den Warenflussdaten eines gastronomischen



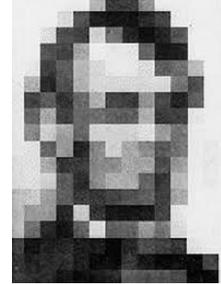
Betriebs kaum Informationen über mögliche Prüffelder oder die Gesamtplausibilität (direkt)

ableiten lassen, offenbart eine **geeignete** Verbildlichung wie in der rechtsseitigen Abb. sowohl eindeutig die strenge Abhängigkeit beider Betriebsfaktoren als auch ein hervorstechendes Einzelprüffeld mit dem relativ zu hohen **roten Wareneinsatzpunkt** im April.

²¹ *Barnard* in einer US-amerikanischen Zeitschrift der Werbebranche 1921: „One look is worth a thousand words“, nachzulesen unter: www.oai.de/de/component/content/article/45-publikationen/sprichwort/811-ein-bild-sagt-mehr-als-tausend-worte.html.

In der Umsetzung der mehrperspektivisch (s. [Pkt. 2.3](#)) visualisierenden SRP wurden zudem **wichtige bestehende Erkenntnisse zur Datenanalyse** berücksichtigt wie insbesondere:

- ☒ Nach den **Fuzzy-Theorien** zur Bedeutung der (Un-)Schärfe ist die Aggregation der Daten entscheidend dafür, welche Informationen erkennbar sind. So lässt sich trotz grober Darstellung im nebenstehenden Bild²² das Antlitz von Abraham Lincoln erkennen. Dieses Wissen ist für die Effizienz der Datenanalyse relevant, denn je verdichteter die zu prüfenden Daten sind, umso schneller und einfacher gestaltet sich die Untersuchung. Auf der anderen Seite wird der Effizienzgedanke von der abnehmenden Effektivität – also der Wahrscheinlichkeit des Erkennens – begrenzt. Zuletzt bewirkt eine zu große Informationsfülle i.d.R. auch eine Überforderung des Prüfungspersonals, so dass zwar Details aber gröbere Regeln und Fehler nicht mehr gesehen werden.
- ☒ Die in der professionellen Datenanalyse²³ präferierten **Koinzidenzanalysen** untersuchen den Einfluss aller möglichen Faktorkombinationen auf Risiken bzw. Chancen und geben die Ergebnisse als komfortabel bedienbare „Rankings“ aus. Mit der (nahezu) vollständigen Verarbeitung der verfügbaren Informationen erfolgt die Datenanalyse vor allem ergebnisoffen und kann somit auch eher unerwartete Fehler und Manipulationen detektieren.
- ☒ Als weltweit verbreitetes Lösungsmodell für komplexe Fragestellungen und Probleme orientiert sich die (biokybernetische) **Sensitivitätsanalyse** von Vester²⁴ an Mustern in der Datenstruktur, weil es „[s]imple Ursache-Wirkungs-Beziehungen [...] nur in der Theorie [gibt], nicht in der Wirklichkeit.“²⁵ Diese Idee eröffnet in der Übertragung auf die Prüfung eine Ausweitung der unter Wirtschaftsprüfern üblichen ABC-Analyse mittels Ergebnisverbildlichung, wodurch wesentlich detaillierter Fehler und Manipulationen aufgedeckt werden können.
- ☒ Im Rahmen der **statistischen Stichprobenverfahren** wie z.B. „Monetary Unit Sampling (MUS)“ wird die Unsicherheit der resultierenden Aussage mit Hilfe des mathematischen Spezialgebiets „Wahrscheinlichkeitslehre“ qualifiziert, weshalb sie gesetz-



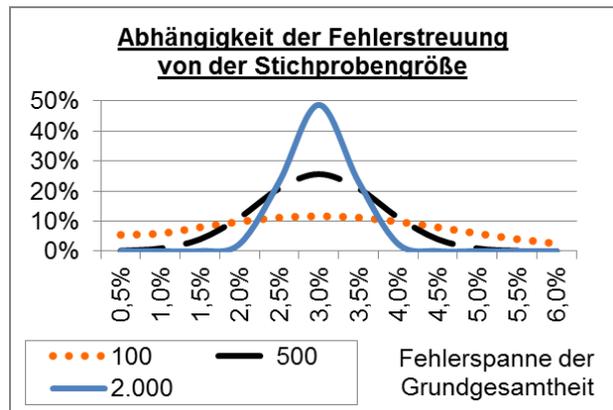
²² Bildquelle: Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 54.

²³ Vgl. z.B.: „HyperCube® [der Fa. BearingPoint] wird vom Pasteur-Institut als das derzeit leistungsstärkste Datenanalyse-Tool auf dem Markt angesehen. [...] ein neuartiger Datenbestimmungsalgorithmus, der die Kombinationen von Variablen identifizieren kann, die eine optimale Erklärung der zugrundeliegenden Fragestellung ermöglichen“, nachzulesen unter: <http://www.bearingpoint.com/de-de/expertise/enabling-technology/hypercube>.

²⁴ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012.

²⁵ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15.

lich zugelassen²⁶ und in der Wirtschaftsprüfung²⁷ verbreitet sind. So gibt diese Abb.²⁸ die Eindeutigkeiten der Übertragung eines 3 %-igen Stichprobenfehlers auf die Grundgesamtheit wieder: Während bei einer **großen Stichprobe von 2.000 Vorgängen** der tatsächliche Gesamtfehler nur geringfügig um die 3 % streut, kann bei einem **kleinen untersuchten Auszug von z.B. 100 Einzelsachverhalten** nicht problemlos auf die Gesamtheit geschlossen werden. Diese konkreten, reproduzier- und damit vergleichbaren Erkenntnisse der Statistik eignen sich besonders für die Einordnung von unsicheren Aussagen, wie sie grundsätzlich vorkommen, wenn die Vergangenheit anhand unvollständiger Dokumente nachvollzogen werden soll.



Für die Ausgestaltung des SRP-Prüfungsduktus und der SRP-Vorlagen, welche die technischen Anwendungsanforderungen möglichst niedrig halten sollen, sind dieses Wissen und die umfangreichen Erfahrungen aus der täglichen Anwendung fundamental. Zudem steht in der Unterrichtung der SRP das umfassende Verständnis der Einzelmethoden im Vordergrund, weil bspw. die Varianten von Zeitreihenanalysen eine stark differierende Eignung für unterschiedliche Prüfungsfragen aufweisen (s. [Pkt. 2.2.](#)).

Damit soll die SRP zur Gleich- und Gesetzmäßigkeit der Besteuerung infolge einer (zeitgemäß) **effektiven Außenprüfung** beitragen, denn „die Finanzbehörden haben alle notwendigen Maßnahmen zu ergreifen, um die entscheidungserheblichen Tatsachen aufzuklären“²⁹. Art und Umfang sollen den „Grundsatz der Verhältnismäßigkeit [...] beachten“³⁰ und werden durch die Mitwirkungspflichten der Steuerpflichtigen beschränkt³¹, bilden aber als **primäre** Aufgabe der Betriebsprüfungsstellen gemäß § 2 BpO die entscheidende Grundlage für die **anschließende** rechtliche Beurteilung – vergleichbar der sorgfältigen Diagnostik in der Medizin als Voraussetzung für die erfolgreiche Therapeutik.

²⁶ § 241 Abs. 1 HGB: „Bei der Aufstellung des Inventars darf der Bestand der Vermögensgegenstände nach Art, Menge und Wert auch mit Hilfe anerkannter mathematisch-statistischer Methoden auf Grund von Stichproben ermittelt werden.“

²⁷ IDW PS 310 „Repräsentative Auswahlverfahren (Stichproben) in der Abschlussprüfung“, nachzulesen unter.: <https://www.idw.de/idw/verlautbarungen/idw-ps-310/43204>.

²⁸ Aus dem Grund der besseren Übersichtlichkeit wird die Fehlerstreuung mit der stetigen Normalverteilung abgebildet, die approximativ den diskreten Verhältnissen der einschlägigen Hypergeometrischen Verteilung entspricht.

²⁹ Anwendungserlass zu § 88 AO Nr 1.

³⁰ Anwendungserlass zu § 88 AO Nr 1.

³¹ Anwendungserlass zu § 88 AO Nr 2.

Entsprechend der Bedeutung der Sachverhaltsaufklärung definiert das Verfahrensrecht mit dem Verifikationsprinzip eine besondere Sphärenverantwortung im **Beweislastverhältnis des § 158 AO**. Dieses gründet sich auf die Einsicht, „dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann“³². Danach führt „eine Verprobung usw, [nach der es] unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt“³³, zum Verlust der Vermutung sachlicher Richtigkeit – bis zur möglichen Konsequenz der ermessensfreien Schätzungspflicht³⁴. Insofern hat eine (zeitgemäß) wirksame Prüfungstechnik zur möglichst weitreichenden Sachverhaltsaufklärung eine grundlegende und direkte Bedeutung für den Schutz der Gleich- und Gesetzmäßigkeit der Besteuerung durch die Außenprüfung.

2.1. Heutige Vollzugssituation und Steuerausfallrisiken

Als Folge der Digitalisierung und Globalisierung der Wirtschaft haben sich die „**Herausforderungen der digitalen Betriebsprüfung**“³⁵ wesentlich erhöht: Rasant anwachsende Mengen an Betriebsdaten, spezielle Fehlerquellen der (weitgehend) automatisierten Datenverarbeitung und die nur schwierig verifizierbare Originalität (Unverändertheit) von elektronischen Aufzeichnungen³⁶ sind Hauptfaktoren für immer größere Probleme beim Auffinden von Fehlern und Manipulationen in Besteuerungsdaten.

„Die heutigen Datenmengen und qualitativen Ansprüche übersteigen zunehmend die Möglichkeit einer manuellen Prüfung. Relevante Aussagen über ein Prüfungsobjekt können heute nicht mehr allein mit der Durchsicht von Listen, simpler Belegprüfung oder durch Stichproben getroffen werden.“³⁷ Vom Prüfungspersonal wird ein „hohes Maß an Methodenwissen, Flexibilität und Kreativität“³⁸ gefordert. So erkennt z.B. auch der BFH: „Elektronische Kassensysteme sind durch Umprogrammierung in nahezu beliebiger Weise manipulierbar;

³² Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5.

³³ Anwendungserlass zu § 158 AO.

³⁴ § 162 Abs. 2 S. 2 AO.

³⁵ *Kulosa (RiBFH)*, „Herausforderungen der digitalen Betriebsprüfung, u.a. bei bargeldintensiven Betrieben“, FR 2017, 501.

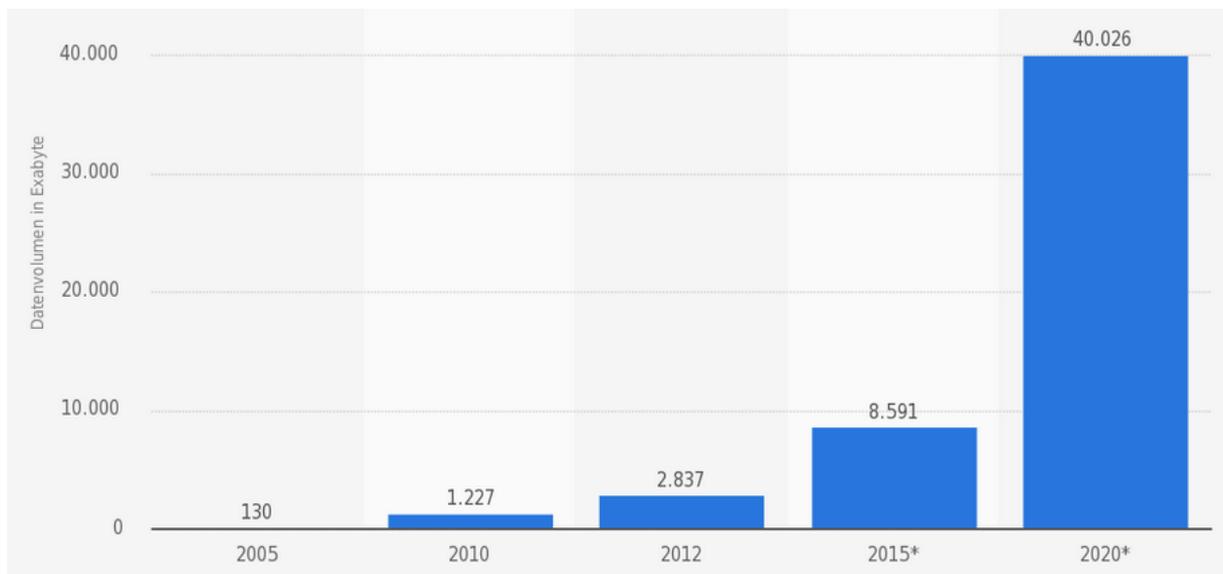
³⁶ *Huber/Reckendorf/Zisky*, „Die Unveränderbarkeit der Kassenbuchführung § 146 Abs. 4 AO im EDV-Zeitalter und INSIKA – Teil 1: Historische Wurzeln und Realzustand“, BBK 12/2013, 567, „Teil 2: Probleme, Ursachen und Lösungen“, BBK 13/2013, 610 sowie „Teil 3: Konzept, Technik, Praxis“ BBK 14/2013, 663.

³⁷ *Ehlers*, „Tax Compliance – Neue Beratungsfelder bei Betreuung kleiner und mittlerer Unternehmen“, NWB 18/2012, 1535, 1540.

³⁸ Internetseite des Lehrstuhls für Wirtschaftsprüfung, Unternehmensrechnung und Controlling von Prof. Mochty an der Universität Duisburg-Essen, nachzulesen unter: <https://www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik>.

von derartigen Manipulationsmöglichkeiten machen Teile der betrieblichen Praxis nach dem Erkenntnisstand des Senats durchaus Gebrauch³⁹. In der Konsequenz sieht der BFH „auf Seiten der Finanzverwaltung ein durchaus nachvollziehbares Bedürfnis [nach] moderne[n] Prüfungsmethoden“⁴⁰. Die (technischen) Anforderungen an das Prüfungspersonal zur Erfüllung der vorrangigen Aufgabe „Sachverhaltsaufklärung“⁴¹ haben sich in nahezu jedem Prüfungsbereich erhöht, wie z.B. auch der umfangreiche Maßnahmenkatalog der OECD⁴² zur Bekämpfung von Steuervermeidungsstrategien international tätiger Unternehmen zeigt.

Auf der anderen Seite eröffnet die digitale Außenprüfung durch die enorme Geschwindigkeit der elektronischen Datenverarbeitung mit der zielführenden Vernetzung geeigneter Prüfungsmethoden (s. [Pkt. 2.3.](#)) erweiterte Möglichkeiten zur weitreichenden Sachaufklärung. Dadurch wird die Bearbeitung von diversen (alltäglichen) Prüffeldern schneller und vollständiger, für einige sogar erstmalig über eine Stichprobenauswahl hinaus durchführbar.



Die **exponentielle Zunahme des Datenvolumens**⁴³ geht darauf zurück, dass „[d]ie Buchführung [...] sich von der Durchschreibebuchführung hin zu einer umfassenden Datenverarbeitung entwickelt [hat], die elektronische Form ist nunmehr die übliche. [...] Diese Entwicklung geschah auf Betreiben und im Interesse der Unternehmen, weil es betriebswirtschaftlich notwendig war und ist, diese Einzeldaten zu erzeugen, z.B. für ein aussagekräftiges Controlling oder zur Arbeitsprozesssteuerung. [...] Dabei steigt die Anzahl elektronisch

³⁹ UrI. v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

⁴⁰ UrI. v. 25.3.2015, BStBl. II 2015, 743, Rz. 54.

⁴¹ § 2 BpO.

⁴² Nachzulesen unter: <http://www.oecd.org/ctp/beps.htm>.

⁴³ Bildquelle: <http://www.informatik.uni-oldenburg.de/~iug16/gv> mit Verweis auf: <https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen>.

überwachter Prozesse stetig.⁴⁴ Da zudem bekannt ist, dass die automatisch generierten Betriebsdaten durchaus nennenswerte Fehlerquellen beinhalten,⁴⁵ wächst die Bedeutung des zielgenauen und zügigen Auffindens von Fehlern und Manipulationen ständig an.

Systemprüfungen in den Groß- und Konzernbetriebsprüfungsstellen haben über die Datenmigration hinaus auch in automatisierten Schlüsselungen, Zusammenführungen und Konsolidierungen von Betriebsdaten mitunter gravierende Unrichtigkeiten offenbart, weshalb bspw. die Untersuchung von Datenschnittstellen als wesentliches neues Prüffeld angesehen wird⁴⁶. Nach kriminalpsychologischen Erkenntnissen kann nur eine hohe Aufdeckungswahrscheinlichkeit von Zuwiderhandlungen in Kombination mit einer angemessenen Sanktionierung die (weitgehende) Einhaltung von (Steuer-)Gesetzen sicherstellen.⁴⁷ Unabhängig von der Art und Größe eines zu prüfenden Unternehmens gewinnt die zeitgemäße Arbeitsweise der Außenprüfung somit eine zunehmende Bedeutung für die Gleich- und Gesetzmäßigkeit der Besteuerung. Die Literatur fordert schon länger eine den heutigen Risiken und Möglichkeiten angemessene Verifikation.⁴⁸

Bereits seit 2003 warnt der BRH vor „[d]rohende[n] Steuerausfälle[n] aufgrund moderner Kassensysteme“⁴⁹. In einem Bericht hat die OECD die Prüfungssituation im Zeitalter **raffii- nierter (digitaler) Manipulationswerkzeuge** mit beispielhaften „Techniken der Umsatzverkürzung, Aufdeckungsstrategien und [s]taatliche[n] Maßnahmen“⁵⁰ zusammengefasst. Das medial⁵¹ dominierende Steuerthema findet ebenfalls regelmäßig in der Literatur⁵² und sogar in Wikipedia⁵³ Beachtung.

⁴⁴ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876 u. 877.

⁴⁵ Vgl. z.B.: Newsletter 4-2016 des Forums „Elektronische Steuerprüfung“ v. 22.4.2016: „(IT-)Management: Ein Drittel aller Unternehmen verlieren Daten bei der Migration: Ein Drittel (32 %) der Organisationen verlieren Daten, während sie Daten zwischen Geräten migrieren oder Betriebssysteme aktualisieren. Das ist das Ergebnis einer weltweiten Befragung unter fast 600 IT-Administratoren“, nachzulesen unter: http://elektronische-steuerpruefung.de/newsarchiv/news_16_04.htm?m=1450705466#Beitrag7.

⁴⁶ *Webel/Danielmeyer*, „Schnittstellenverprobung elektronischer Betriebsverwaltungen – das kommende Standardprüffeld?“, StBp 2015, 353.

⁴⁷ Vgl. z.B.: *Smettan*, „Kriminelle Bereicherung in Abhängigkeit von Gewinn, Risiken, Strafen und Moral“, Kriminologische Forschungsreihe aus dem Max-Planck-Institut für ausländisches und internationales Strafrecht Bd. 63, 1992.

⁴⁸ *Fischer (Vors. RiBFH a.D.)*, „Das EuGH-Urteil Persche zu Auslandsspenden – die Entstaatlichung des Steuerstaates geht weiter“, FR 2009, 249.

⁴⁹ Bundesrechnungshof, „Bemerkungen 2003 zur Haushalts- und Wirtschaftsführung des Bundes“, Pkt. 54.

⁵⁰ OECD-Bericht „Umsatzverkürzung mittels elektronischer Kassensysteme – eine Bedrohung für die Steuereinnahmen“, 2013, nachzulesen unter: https://www.oecd.org/ctp/crime/Sales_suppression_German_website.pdf.

⁵¹ Vgl. z.B.: Zeit online v. 20.2.2014, „Zu Gast bei Betrügern – Manipulierte Kassen, falsche Rechnungen und Schummelsoftware: Mit welchen Tricks Gastronomen Steuern hinterziehen – und wie ihnen Fahnder auf die Schliche kommen“, nachzulesen unter: <https://www.zeit.de/2014/09/steuerhinterziehung-gastronomie>.

⁵² S. z.B.: *Huber*, „Gedanken zur Zukunft des steuerlichen Risikomanagements im Erlösbereich“, StBp 2014, 121 u.w., *Pump*, „Rechtsfolgen bei Verwendung von Manipulationssoftware (Zappersoftware) – Steuerrecht,

Die systemimmanenten oder als Zusatzmodul käuflichen Eingriffsfunktionen nutzen in den (komplexen) Softwarelösungen für die Kasse(n), Warenwirtschaft, Rechnungslegung,

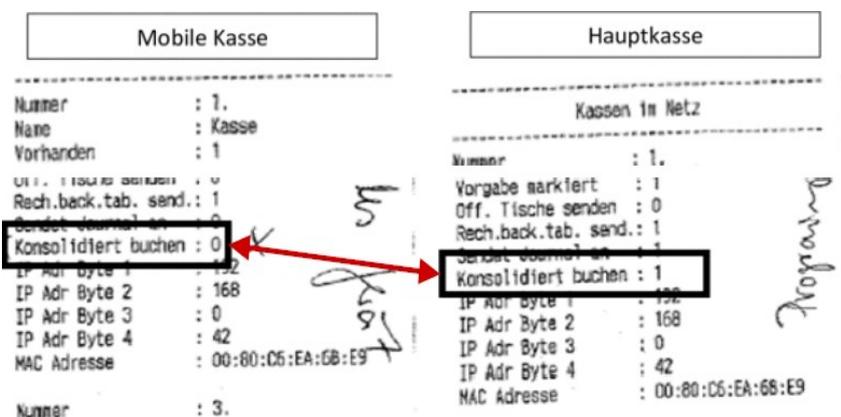
Diese Software unterstützt z.B. die fortlaufenden Rechnungsnummern. Um bei Bearbeitungsfehlern jedoch Korrekturen vornehmen zu können, ist das das Ändern von Rechnungsnummern jederzeit möglich! ?

Genauso ist es möglich jederzeit Daten unprotokolliert (also restlos) zu löschen ?

Bevor Sie diese Software benutzen, muss sich der Anwender beim Finanzamt, beim Steuerberater oder bei anderen zuverlässigen Quellen über diese Regelungen und Vorschriften informieren. !

Kundenverwaltung etc. der Unternehmen die Intransparenz der Systeme und die leichte Veränderbarkeit digital gespeicherter Informationen aus. Dabei wird z.T. offensiv mit den Möglichkeiten der Steuerhinterziehung ge-

worben.⁵⁴ Einen Überblick über die weitreichenden Manipulationsoptionen im heutzutage



üblichen betrieblichen e-Management findet sich bei *Teutemacher* mit Beispielen aus der Fahndungspraxis wie in nebenstehendem Auszug zur Systemeinstellung: „Es wird so der Anschein erweckt, als

ob alle Umsätze im Kassensystem erfasst werden. Dies ist jedoch ein Trugschluss, wie die Programmierung des Mobilgeräts [...] zeigt, denn dort wurde bei ‚konsolidiert buchen‘ eine ‚0‘ programmiert. Dies führt dazu, dass Umsätze, die über dieses Kassenterminal erfasst wurden, nicht im Tagesumsatz dokumentiert werden.“⁵⁵ Der mögliche Umfang von Steuerhinterziehungen mit Hilfe digitaler Aufzeichnungssysteme lässt sich aus einem Beschluss des FG Rheinland-Pfalz erahnen, nach welchem die hinterzogenen Steuern eines einzelnen Eiscafés für den erweiterten Prüfungszeitraum knapp € 2 Mio. betragen.⁵⁶

Steuerstrafrecht und besonderes Strafrecht“, DStZ 2013 Nr. 9, 299, *Huber*, „Steueraufsicht und Betriebsprüfung in der Zeit der Kassenandroiden und ohne INSIKA – Grundgedanken, Ziele, Risiken und zweitbeste Lösungen“, StBp 2012, 301 u.w. sowie *Härtl/Schieder*, „Ordnungsmäßigkeit digital geführter Erlösaufzeichnungen – Elektronische Registrierkassen und digitale Erlöserfassungssysteme im Brennpunkt des Steuerrisikos Erlöserfassung“, StBp 2011, 33 u.w.

⁵³ Nachzulesen unter: [https://de.wikipedia.org/wiki/Zapper_\(Software\)](https://de.wikipedia.org/wiki/Zapper_(Software)).

⁵⁴ *Huber*, „Gedanken zur Zukunft des steuerlichen Risikomanagements im Erlösbereich – Teil II“, StBp 2014, 153, 155 mit der Abb. und 156 mit einem Zitat aus der dortigen Abb. 10 „Werbung für Handwerkersoftware“: „Insbesondere von Vorteil ist, dass man jederzeit alle Einträge rückgängig machen oder ändern kann. Da speziell im Handwerk einiges anders läuft (z.B. Schwarzgeld) ist dies auch für Sie wichtig, denn Umbuchungen oder Löschungen sind nachträglich nicht mehr sichtbar.“

⁵⁵ *Teutemacher*, „Handbuch zur Kassenführung – Praxishandbuch für die rechtssichere Umsetzung“, NWB Verlag, 2015, 152 mit der Abb. und dem Zitat.

⁵⁶ Beschl. v. 7.1.2015 – 5 V 2068/14; zitiert von: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

Besonders im Bereich der digitalen Erfassung von Bareinnahmen wirkt sich der Wettlauf von Verbrechen und Aufdeckung insofern aus, dass das Prüfungspersonal zur effektiven Durchführung der Arbeit über technische Kenntnisse sowie ein „hohes Maß an Methodenwissen, Flexibilität und Kreativität“⁵⁷ verfügen muss. Dabei kommt erschwerend der Eindruck hinzu, dass einige VertreterInnen der steuerberatenden Berufe ihrer Vorprüfungsfunktion als Organ der Rechtspflege⁵⁸ – z.B. im Hinblick auf elektronische Vorsysteme – nicht (mehr) ausreichend nachkommen.⁵⁹

2.2. Moderne Prüfungsmethodik in der digitalisierten Welt

Schon die **Formalprüfung** zur Feststellung, ob Besteuerungsunterlagen infolge ihrer Ordnungsmäßigkeit eine inhaltliche Prüfung (in angemessener Zeit) zulassen, stellt in Verbindung mit elektronischer Aufzeichnung und Aufbewahrung höhere Anforderungen an das Prüfungspersonal. Während „[d]er sachverständige Dritte [...] über seine beruflich erlernten Buchhaltungskennntnisse hinaus in der Papier-/Tinte-/Bleistift-/Radierer-Zeit keine außergewöhnlich hohe technische Qualifikation [brauchte], um die Erfüllung der gesetzlichen Vorgaben feststellen zu können“⁶⁰, müssen heutzutage komplexe Betriebsverwaltungsprozesse verstanden werden⁶¹. Nur dann ist eine „teleologisch[e]“⁶² Interpretation der Ordnungsmaßnahmen – dem Gesetzeszweck folgend – möglich.

Die Rechtsprechung hat zwischenzeitlich in diversen Verfahren auf die grundlegend veränderten Bedingungen der digitalen Ordnungsmäßigkeitsprüfung reagiert, die gem. der Normenkette § 158 i.V. mit § 162 Abs. 2 S. 2 AO direkt bis zur Schätzung der Besteuerungsgrundlagen führen kann⁶³: Im Allgemeinen sind die Ordnungskriterien „**funktional**“⁶⁴ auf

⁵⁷ Internetseite des Lehrstuhls für Wirtschaftsprüfung, Unternehmensrechnung und Controlling von Prof. Mochty an der Universität Duisburg-Essen, nachzulesen unter: <https://www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik>.

⁵⁸ § 57 StBerG.

⁵⁹ *Pump/Heidl*, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenbuchführung – Teil II“, StBp 2014, 162 u. 204, 163, 206: „Unverantwortliche Sorglosigkeit kann die Berater dazu verleiten, sich nicht um die Kassenführung ihrer Mandanten zu kümmern. Das wird damit begründet, dass dafür kein Prüfungsauftrag bestehe. [...] Die Berufspflichten für Berater sind streng. [...] Das Prinzip ‚Augen zu und ab mit der falschen Steuererklärung‘ ist töricht. Das würde die vom Berater zu erwartende Klugheit widerlegen, weil er sich durch das Mandat strafbar machen würde.“

⁶⁰ *Huber/Reckendorf/Zisky*, „Die Unveränderbarkeit der Kassenbuchführung § 146 Abs. 4 AO im EDV-Zeitalter und INSIKA – Teil 1: Historische Wurzeln und Realzustand“, BBK 12/2013, 567, 572.

⁶¹ Vgl. insb.: „Grundsätze zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)“, BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450.

⁶² *Tipke/Kruse*, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 12.

⁶³ Vgl. z.B.: *Nöcker (RiBFH)*, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3550: „Liegen keine materiellen Buchführungsmängel vor, entscheidet die Gewichtigkeit der formellen Buchführungsmängel darüber, ob eine Schätzungsbefugnis des Finanzamts vorliegt.“

den Einzelfall anzuwenden, so dass im Besonderen bspw. das Fehlen der Systemeinstellungen einer elektronischen Kasse einen gravierenden Buchführungsmangel darstellt⁶⁵. Damit „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO)⁶⁶, sind eine einzelfallbezogen fundierte Formalprüfung und eine Auslegung von Ordnungsmängeln gem. des Verifikationsprinzips⁶⁷ Voraussetzung.

Seit Einführung der digitalen Außenprüfung zum 1.1.2002 werden regelmäßig neue bzw. erweiterte Methoden für die zeitgemäße **materielle Prüfung** in Veröffentlichungen⁶⁸ vorgestellt. Durch den Vorsprung der digitalen Prüfung, große Datenmengen in sehr hoher Geschwindigkeit analysieren zu können, sollen die infolge der elektronischen Betriebsverwaltung gestiegenen Anforderungen (s. [Pkt. 2.1.](#)) besser bewältigt werden. Dabei haben auch die Medien⁶⁹ und professionellen Anbieter⁷⁰ den wahrnehmungsphysiologischen Vorteil der Datenverbildlichung (s. [Pkt. 2.](#)) erkannt.

Wesentlicher Trend der Entwicklung in der Wirtschaft ist hierbei eine hochgradige Automatisierung der Prozesse: „Die an Systemschnittstellen erforderlichen manuellen Tätigkeiten werden immer mehr durch den Einsatz von Robotic Process Automation ersetzt, Entscheidungsprozesse werden dabei durch den Einsatz programmierter, selbstlernender Entscheidungsheuristiken, die auf Basis künstlicher Intelligenz arbeiten, unterstützt.“⁷¹ Ebenso zeichnet sich unter Begriffen wie z.B. „Dashboards“⁷² das neue Analysemantra „viel hilft viel“ ab,

⁶⁴ FG Hamburg v. 26.8.2016 – 6 V 81/16, 1. Leitsatz.

⁶⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, 5. Leitsatz.

⁶⁶ BFH v. 15.2.1989, BStBl. II 1989, 462: „[...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

⁶⁷ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

⁶⁸ Vgl. z.B.: *Sosna*, „Einsatz statistischer Methoden zur Risikoanalyse, Recherche und Lokalisierung von Steuerausfällen“, StBp 2000, 41 u. 68, *Huber*, „Weiterentwickelte und neue Methoden der Überprüfung, Verprobung und Schätzung“, StBp 2002, 199, 233, 258 u. 293, *Gebbers*, „Die Strukturanalyse“, StBp 2007, 225, 257 u. 289, *Posch*, „Digitale Ziffernanalyse“, StBp 2010, 338, *Hentrich/Pyrcek*, „Compliance und Fraud Monitoring im Zeitalter von digitaler Transformation und Big Data“, BB 2016, 1451, u.w.

⁶⁹ Bspw. Wirtschaftswoche online v. 21.5.2015, nachzulesen unter:

<https://www.wiwo.de/politik/deutschland/fiskus-4-0-mit-big-data-auf-steuerfang/11770850.html>.

⁷⁰ Vgl. z.B.: Internetblog der Fa. Audicon zu IDEA V10: „Grafische Darstellungen lassen sich im Allgemeinen leichter erfassen als die Darstellung der Daten in Tabellenform. Daher ist ein besonderes Feature des neuen Releases die Visualisierung“, nachzulesen unter: <https://blog.audicon.net/audit-analytics/ideav10-visualisierung-und-ermitteln.html> oder Visualisierungsprogramm „Visual Analytics“ des amerikanischen, weltweit operierenden Softwarehauses Statistical Analysis Systems, nachzulesen unter:

http://www.sas.com/de_de/software/business-intelligence/visual-analytics.html.

⁷¹ *Kompenhans/Wermelt*, „Big Data“ – Potenziale für innovative Abschlussprüfungen“, BB 2018, 299.

⁷² Nachzulesen unter: <https://blog.audicon.net/audit-analytics/ideav10-visualisierung-und-ermitteln.html>.

indem auf Knopfdruck eine Vielzahl von Visualisierungen eingepflegter Daten ausgegeben werden.

Die **SRP** verfolgt bei der Umsetzung der digitalen Außenprüfung eine andere Strategie: In ihrem Zentrum steht das **Verständnis der einzelnen Prüfungsmethoden**. Die Fortbildungen und Hilfsmittel – insbesondere die SRP-Vorlagen – sollen die technische Umsetzung durch das Prüfungspersonal erleichtern und beschleunigen, **ohne den Prüfungsprozess zu übernehmen**. Der Grund für diese Fokussierung liegt in den folgenden zwei Überlegungen:

- ! Viele Datenanalysemethoden sind für bestimmte Fragestellungen sehr gut geeignet, für andere hingegen nicht (*s.a. das Bsp. zum folgenden Abschnitt „Methodenwissen“*). Um die Gefahr von Ergebnisinterpretationen auf Basis ungeeigneter Analyseansätze zu minimieren, werden Methodenunterschiede geschult, zielführende Kombinationen anhand des SRP-Prüfungsduktes „vom groben Überblick zu den Details“ vermittelt und in den SRP-Vorlagen technisch so zur Verfügung gestellt, dass nur geringe IT-Kenntnisse erforderlich sind und die Durchführung dennoch in der Hand des Prüfungspersonals verbleibt.
- ! Während die – vom Menschen entwickelte – Datenanalyse auf der Grundlage von Ordnung funktioniert, entscheidet das menschliche Gehirn chaotisch-assoziativ.⁷³ Somit verbindet der optimale Einsatz der digitalen Datenverarbeitung diese zwei sehr unterschiedlichen Analyseformen. Demgegenüber würde die Konzentration auf nur eine Prüfungsressource „Mensch“ oder „Computer“ den Verzicht auf wichtiges Prüfungspotential bedeuten.

Am Anfang jeder (zielgerichteten) Datenanalyse steht eine zu klärende Frage bzw. zu lösende Aufgabe. Für die steuerliche Außenprüfung ergibt sich diese aus den gesetzlichen Aufträgen der Plausibilitätsprüfung des Gesamtbilds⁷⁴ von (Teilen der) Besteuerungsaufzeichnungen gem. § 158 AO sowie der effizienten⁷⁵ Aufdeckung von betriebsspezifischen Einzelprüffeldern.

Im nächsten Schritt sollten die in Betracht kommenden Werkzeuge zur Beantwortung der Prüfungsfrage auf ihre jeweilige Eignung hinterfragt werden, weshalb *Mochty* mit einem

⁷³ Intuitiv scheint dieser Unterschied den Führungskräften der Wirtschaft bewusst zu sein: „Manager vertrauen mehr auf ihr Bauchgefühl als auf B[usiness]I[n]telligence“ zitiert CRN online v. 24.2.2010 eine entsprechende Studie.

⁷⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

⁷⁵ § 7 BpO: „Prüfungsgrundsätze: Die Außenprüfung ist auf das Wesentliche abzustellen. Ihre Dauer ist auf das notwendige Maß zu beschränken. Sie hat sich in erster Linie auf solche Sachverhalte zu erstrecken, die zu endgültigen Steuerausfällen oder Steuererstattungen oder -vergütungen oder zu nicht unbedeutenden Gewinnverlagerungen führen können.“

„hohe[n] Maß an **Methodenwissen**, Flexibilität und Kreativität“⁷⁶ große Anforderungen an das Prüfungspersonal erkennt. Auch *Ehlers* erwartet permanente „Maßnahmen der Qualitätssicherung“⁷⁷. Theoretische Überlegungen und praktische Erfahrungen sollten bei der Eignungsbewertung von Prüfungsansätzen zusammenkommen, wie es in der wissenschaftlichen Validierung üblich ist. In Rechtsprechung⁷⁸ und Literatur⁷⁹ finden sich zwar z.B. zur Zeitreihenanalyse umfangreiche theoretische Erörterungen, die jedoch oftmals von der inzwischen sehr umfangreichen Praxis⁸⁰ widerlegt werden.

„Prinzipiell sind zahlreiche Varianten eines Zeitreihenvergleichs denkbar“⁸¹, erkennt der BFH in seinem Zeitreihenurteil, ohne die Unterschiede zu ergründen, und lässt sich dennoch zu einem Generalurteil über alle Zeitreihenanalysen mit strengen Anwendungsbeschränkungen hinreißen⁸². Demgegenüber nimmt das FG Hamburg zur unterschiedlichen Ausrichtung von Zeitreihenvarianten konkret Stellung: „Mit Hilfe der doppelskalierten Monatszeitreihe hat die Betriebsprüfung die Daten des Wareneinkaufs den Umsatzerlösen in zwei Einzelgrafiken gegenüber gestellt. Bei dieser Methode tritt das zahlenmäßige Verhältnis zu Gunsten der Abhängigkeit voneinander in den Hintergrund. Wesentlich ist die Reaktion der Faktoren aufeinander. Aus kaufmännischer Sicht wird in einem Unternehmen auf jede Veränderung der Nachfragesituation (annähernd) adäquat mit einer Erhöhung bzw. Absenkung der Einkaufssumme reagiert“⁸³.

In den folgenden Abb. werden die Warenflussverhältnisse aus **Einkauf** und Umsatz eines gastronomischen Betriebs mit Einnahmen-Überschuss-Rechnung gem. § 4 Abs. 3 EStG durch zwei Typen von Monatszeitreihen wiedergegeben. Losgelöst davon, dass diese Fallkonstellation mit länger haltbaren Getränken und Aufzeichnung nach Bezahlung lt. finanzge-

⁷⁶ Internetseite des Lehrstuhls für Wirtschaftsprüfung, Unternehmensrechnung und Controlling von Prof. Mochty an der Universität Duisburg-Essen, nachzulesen unter: <https://www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik>.

⁷⁷ *Ehlers*, „Tax Compliance – Neue Beratungsfelder bei Betreuung kleiner und mittlerer Unternehmen“, NWB 18/2012, 1535, 1541.

⁷⁸ Vgl. z.B. die Kritik von *Wolenski* in „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, Stbg 6/2016, 268 (mit einer beispielhaften Verdeutlichung auf Seite 269 und in Fußnote 5), an den Warnungen des BFH vor möglichen „Hebelwirkungen“, weil die dazugehörigen Rechenbeispiele auf die besondere Zeitreihenvariante und den Einzelfall begrenzt sind (Urt. v. 25.3.2015, BStBl. II 2015, 743, Rz. 44 ff.).

⁷⁹ S. z.B. die erheblichen Sorgfaltsanforderungen an den Einsatz von Zeitreihenanalysen in *Wiggen*, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, die regelmäßig unreflektiert von den Finanzgerichten übernommen werden wie z.B. in BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 55 ff.

⁸⁰ Aus dem (Flächen-)Einsatz der SRP seit 2008 in inzwischen drei Viertel der deutschen Bundesländer durch zurzeit mehr als 2.000 PrüferInnen mit vielen tausend Fällen sind die eklatanten Unterschiede einzelner Varianten der Zeitreihenanalyse bestens bekannt, so dass eine generelle Eignungsbeurteilung ohne praktischen Hintergrund fahrlässig ist, wie das Beispiel des Folgeabsatzes eindrücklich dokumentiert.

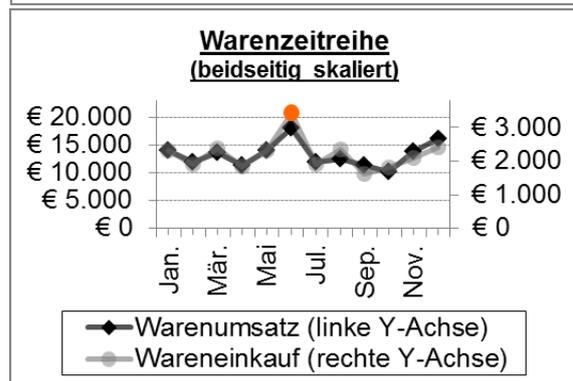
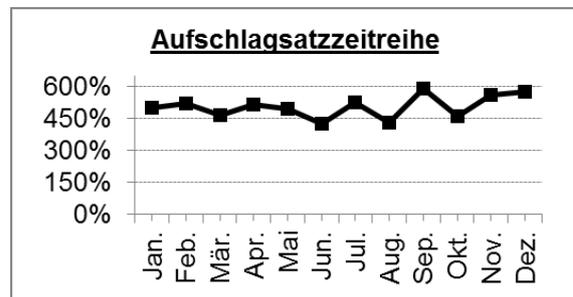
⁸¹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 37.

⁸² BFH v. 25.3.2015, BStBl. II 2015, 743, 1. Leitsatz.

⁸³ Beschl. v. 31.10.2016 – 2 V 202/16, Rz. 48.

richtlicher Einschätzung nicht für Zeitreihenanalysen in Betracht kommt, offenbaren die beiden Diagramme vor allem extrem abweichende Ergebnisse:

1. Die monatliche Aufschlagsatzentwicklung zeigt einen größeren Schwankungsbereich von unter 450 % bis fast 600 %. Diese Beobachtung könnte als implausibles Unternehmerverhalten i.S.d. § 158 AO gedeutet werden.
2. Die mittels Doppelskalierung größenangepasste Gegenüberstellung derselben Ausgangsdaten als Monatssummen ist von gut erkennbarer Abhängigkeit des Warennachkaufs zum Umsatz geprägt. Zudem präsentiert die Datenverbildlichung z.B. mit dem **relativ zu hohen Einkauf im April** ein konkretes Einzelprüffeld



Als analytische Gegenkontrolle bestätigt das statistische Maß für die Abhängigkeit zweier Datenreihen „Korrelationskoeffizient“⁸⁴ mit einem hervorragenden Wert von +0,90 das untere Resultat der **zentralen SRP-Zeitreihe für Plausibilitätsfragen**, während die obere Variante einer Aufschlagsatzzeitreihe⁸⁵ dem nicht entspricht.

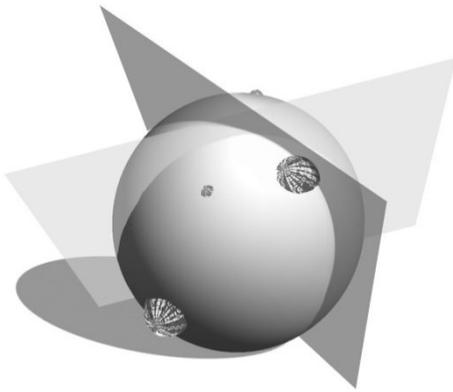
Wegen dieser fundamentalen Unterschiede in der Eignung von Methodenvarianten, welche bisher weitgehend als einheitlich angesehen werden, liegt der Fokus der SRP auf einem ausreichenden Methodenverständnis, damit zielführende Datenvisualisierungen zu einer bestimmten Fragestellung verwendet werden.

2.3. Vorteile mehrperspektivischer Prüfungsnetze wie SRP

Dass **mehrperspektivische Betrachtungen** komplexer Objekte bzw. Situationen erheblich mehr Transparenz herstellen und damit eine wesentlich bessere Beurteilungsgrundlage liefern, ist z.B. aus der Medizin hinlänglich bekannt. Einzelne Schichtaufnahmen von Körperteilen können nur Ausschnitte der Information erfassen. So wird durch die zwei

⁸⁴ Als lineares Abhängigkeitsmaß nimmt der Korrelationskoeffizient Größen von „-1“ für die 100 %-ige Gegenläufigkeit der Grafen über „0“ bei nicht vorhandener Beeinflussung bis zu „+1“ für eine vollständig (positive) Gleichläufigkeit ein.

⁸⁵ Auch dem BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743) liegt mit der „Gleitschlittenzeitreihe“ [Nöcker (RiBFH), „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548] eine Form der Aufschlagsatzzeitreihe zu Grunde.



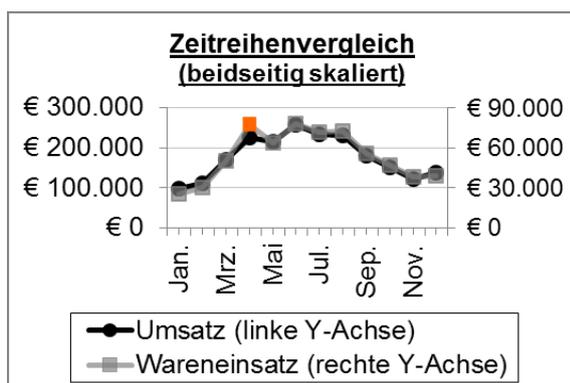
Schichtbilder der Kugel – als Stellvertreter für ein komplexes Objekt – in der Abb.⁸⁶ lediglich eine von vier Auffälligkeiten entdeckt. Seit die technischen Möglichkeiten vorhanden sind, wird deshalb in der Diagnostik und Operationsbegleitung mit vielen Schichtaufnahmen gearbeitet, die computergestützt zu einer dreidimensionalen Rekonstruktion hochgerechnet werden.

Auch in der Datenanalyse führen mehrere *geeignete* Verbildlichungen als **Prüfungsnetz** zu einem besseren Verständnis der Informationen, weil jede

- Sortierung,
- Filterung,
- Aufsummierung bzw. anderweitige statistische Zusammenfassung⁸⁷,
- regelbasierte Verbildlichung per Zeitreihenanalytik,
- musterbasierte Häufigkeitsauszählung nach differenzierbaren Merkmalen wie z.B. dem USt-Schlüssel usw.

mit der jeweiligen Fokussierung weitere Inhalte der Datenmenge offenlegt, so dass die Zusammenführung als Gesamtbild eine viel vollständigere Beurteilungsgrundlage bilden muss als bspw. das Ergebnis eines einzelnen Prüfmakros.

Durch das zielführende Verknüpfen mehrerer regel- und musterbasierter Visualisierungen entlang des SRP-Prüfungsduktus „vom groben Überblick zu den Details“ gelingt das wirksame und zügige Auffinden von betriebsspezifischen Prüffeldern auch in großen Umfän-



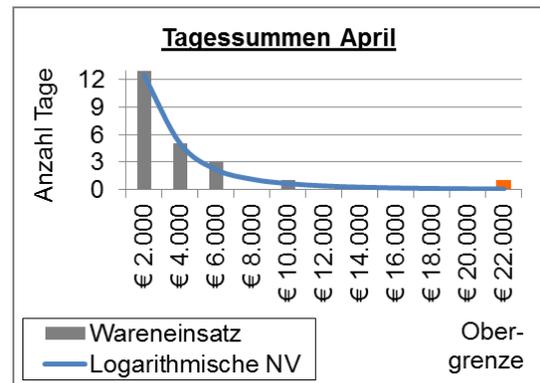
gen von Steuerungsdaten (s. [Pkt. 5.2.](#)). Dabei beleuchtet die regelbasierte Zeitreihenanalytik der SRP-Vorlage BWA die Entwicklungen und Abhängigkeit von Periodensummen zweier Betriebsfaktoren wie in der beispielhaften Abb. mit dem monatlichen Warenfluss eines gastronomischen Betriebs in doppelskaliert größenangepassten Einzelgrafiken (s.a. [Pkt. 2.2.](#)). Zu

dem hierdurch **auffallenden Monat „April“** ergänzt die nachfolgende musterbasierte Grafik der SRP-Vorlage Stochastik eine wertgestaffelte Häufigkeitsauszählung der täglichen Warenein-

⁸⁶ Bildquelle: Mehret/Wähner, „Prüfungsnetze vs. Einzelmethoden: Ein wichtiger, (bisher) verkannter Vorteil der Summarischen Risikoprüfung (SRP)“, DStR 2018, 314.

⁸⁷ Dem Zweck der beschreibenden (deskriptiven) Statistik folgend wird beim Aufsummieren bzw. Ermitteln von Kennzahlen die Gesamtinformation reduziert, um bestimmte Eigenschaften der Datenmenge herauszuarbeiten.

satzsummen in Gegenüberstellung mit der **spezi-
fischen – aus diesen Daten abgeleiteten – Loga-
rithmischen Normalverteilung** (s. [Pkt. 4.3.4.](#)). Mit
der zielführenden Kombination der beiden Verbild-
lichungsansätze lassen sich verursachende Sach-
verhalte als risikoorientierte Prüffelder effektiv und
effizient auf übersichtliche Teilbestände der Daten
eingrenzen: **ein einzelner Tag mit außergewöhn-**



lich hohem Wareneinsatz. Die SRP-Vorlagen unterstützen diesen Prozess mit ihrem Aufbau,
der dem Prüfungsduktus folgt, und den zur Verfügung gestellten Werkzeugen.

Im Rahmen von **Gesamtbildprüfungen** zu (Teilen von) Besteuerungsdaten ist der posi-
tive Effekt mehrperspektivischer Betrachtung besonders groß. Wohingegen Einzelvorgänge
mit ordnungsgemäßen Unterlagen i.d.R. weitgehend als Ursache-Wirkungs-Beziehungen
wiederhergestellt werden können, gibt es diese einfachen Zusammenhänge bei komplexen
Sachverhalten nicht⁸⁸. So können z.B. bei einer Kalkulation Einzelfaktoren wie der Waren-
verderb oder partielle Sonderpreise das Gesamtergebnis empfindlich verändern.⁸⁹ Entspre-
chend müssen bei Gesamtbildbeurteilungen Aussagen grds. unter erheblich größerer Unsi-
cherheit getroffen werden.

Die Nutzung der digitalen Performance-Vorteile für eine systematisch mehrperspektivi-
sche Betrachtung eines komplexen Prüffelds kann die Aussagegewissheit zur Plausibilität
wesentlich verbessern, indem festgestellte Auffälligkeiten mit vertretbarem Aufwand hinsicht-
lich möglicher Besonderheiten des Einzelfalls hinterfragt werden. Das wirksame und zügige
Auffinden außergewöhnlicher Sachverhalte (s.a. *das Bsp. im Abschnitt „Prüfungsnetz“ ober-*
halb) ermöglicht es, diese Ausnahmeverhältnisse in die Schlussfolgerungen angemessen
einzubeziehen. Dabei hilft insbesondere die Interaktivität (s. [Pkt. 4.2.9.](#)), durch welche bspw.

- Untersuchungen weiter eingegrenzt,
- Einzeldaten nachvollziehbar korrigiert,
- Konjunktoreinflüsse mittels Preisindizes bereinigt oder
- ganze Datenreihen im Datum verschoben werden können, um z.B. den Einfluss der Er-
fassung nach Bezahlung – anstatt nach Durchführung – i.R. der Gewinnermittlung nach
§ 4 Abs. 3 EStG zu simulieren.

⁸⁸ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15: „Unser Dilemma im Umgang mit der Komplexität
unserer Welt lässt sich darauf zurückführen, dass wir wohl darin ausgebildet wurden einfache logische
Schlüsse zu ziehen und nahe liegende Ursache-Wirkungs-Beziehungen zu definieren. [...] Simple Ursache-
Wirkungs-Beziehungen gibt es nur in der Theorie, nicht in der Wirklichkeit.“

⁸⁹ Vom BFH im Urt. v. 25.3.2015, BStBl. II 2015, 743, Rz. 44 ff., als „Hebelwirkungen“ im Zusammenhang mit
Zuordnungsunschärfen in Zeitreihenanalysen bezeichnet.

Nur durch diese Simulationsoptionen kann die aus § 158 AO abgeleitete gemeinsame Sachaufklärungsverantwortung⁹⁰ im Prüfungsprozess weitreichend und zeitsparend umgesetzt werden.

Neben der technischen Unterstützung liegt der Sinn der SRP-Vorlagen darin, das Risiko von Analysefehlern zu verringern und die systematisch mehrperspektivische Prüfungsweise – soweit möglich bzw. sinnvoll – zu standardisieren und über ihre Ergebnisausgaben eine gute Nachvollziehbarkeit⁹¹ zu gewährleisten.

Für die nach umfangreichen Praxiserfahrungen⁹² typischen Gesamtbildergebnisse einer SRP mit redundanten⁹³ Auffälligkeiten bzgl. Ordnungsmäßigkeit und inhaltlicher Plausibilität (s.a. [Pkt. 4.4.](#) mit einem Bsp.) gilt zudem das beweisverstärkende Prinzip der Indiziengesamtheiten⁹⁴.

⁹⁰ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen“.

⁹¹ Im Sinn „des Anspruchs der Steuerpflichtigen auf einen effektiven Rechtsschutz durch die Gerichte (Art. 19 Abs. 4, Art. 20 Abs. 3 GG)“ (BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 49).

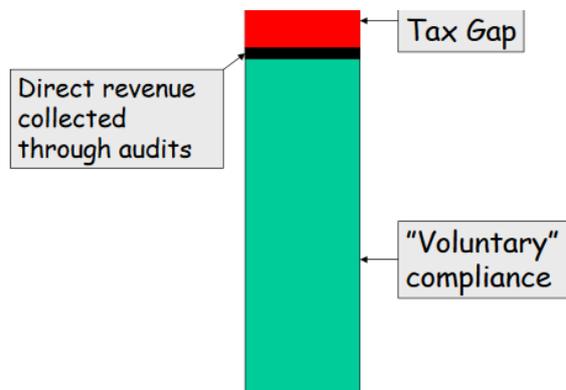
⁹² Vgl. a.: Landesrechnungshof Schleswig Holstein v. 6.2.2012, „Prüfungsmitteilung – Ergebnis der Prüfung ‚Organisation und Arbeitsweise der Betriebsprüfungsstellen‘“, LRH 13 – Pr 1667/2011, Pkt. 5.3.

⁹³ Vgl. z.B.: Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist“ oder *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25: „In der Praxis wird sich die Frage, ob ein Zeitreihenvergleich allein die Beweisvermutung des § 158 AO widerlegt, [...] dadurch relativieren, dass die [...] aufgedeckten Unregelmäßigkeiten regelmäßig mit Unregelmäßigkeiten in der Kassenbuchführung einhergehen.“

⁹⁴ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

3. Einbindung der SRP in Prüfung und Risikomanagement

Ein wirksames **Risikomanagement** der Finanzverwaltung zur Sicherstellung der Gesetz- und Gleichmäßigkeit der Besteuerung muss bei knapp 8 Mio. registrierten Unternehmen und begrenzten Prüfungsressourcen⁹⁵ möglichst frühzeitig ansetzen, denn die Durchsetzung von (Steuer-)Gesetzen verlangt eine erfolgreiche Aufdeckung und spürbare Sanktionierung von Zuwiderhandlungen⁹⁶. Gem. Erkenntnissen der Europäischen Union werden die allermeisten Steuern freiwillig erichtet, wohingegen durch Prüfungsmaßnahmen lediglich ein Anteil im untersten Prozentbereich sichergestellt wird.⁹⁷ Deshalb kommt der Präventivwirkung verbreiteter, regelmäßiger und wirksamer Kontrollen eine überragende Bedeutung zu.



Die SRP setzt an weitgehend unverdichteten Daten an, denn das Aufsummieren zu Periodenwerten vernichtet viele Spuren von Fehlern und Manipulationen (*s.a. das Stichwort „Fuzzy-Theorien“ unter Pkt. 2. und das Zeitreihenbeispiel hierzu unter Pkt. 4.2.2.*). Somit ist eine effektive Umsetzung in Deutschland nach dem derzeitigen Stand des Verfahrensrechts⁹⁸ erst i.R. von Außenprüfungen möglich.

Zur **Einführung der SRP** sind jeweils einwöchige Schulungen zur vorlagengestützten und freien Anwendung Voraussetzung. Hiermit werden die theoretischen Methodengrundlagen

⁹⁵ Laut Monatsbericht des BMF für den November 2017 mit der Betriebsprüfungsstatistik 2016 waren 7.816.301 Betriebe in der Betriebskartei der Finanzämter erfasst und sollten von bundesweit 13.746 PrüferInnen – also mehr als 560 Betriebe pro PrüferIn – intensiv geprüft werden; nachzulesen unter: <https://www.bundesfinanzministerium.de/Monatsberichte/2017/11/Inhalte/Kapitel-3-Analysen/3-7-Ergebnisse-der-steuerlichen-Betriebspruefung-2016.html>.

⁹⁶ Vgl. z.B.: *Smettan*, „Kriminelle Bereicherung in Abhängigkeit von Gewinn, Risiken, Strafen und Moral“, Kriminologische Forschungsreihe aus dem Max-Planck-Institut für ausländisches und internationales Strafrecht Bd. 63, 1992.

⁹⁷ European Commission, Directorate-General, Taxation and Customs Union, Indirect Taxation and Tax administration, Administrative cooperation and fight against fiscal fraud, „Risk Management Guide for Tax Administrations – Fiscalis Risk Analysis Project Group“, (Financial Code: FPG/11), Version 1.02 (February 2006), 60, nachzulesen unter: https://ec.europa.eu/taxation_customs/sites/taxation/files/resources/documents/taxation/tax_cooperation/general/overview/risk_management_guide_for_tax_administrations_en.pdf.

⁹⁸ § 147 Abs. 6 AO: „Sind die Unterlagen nach Absatz 1 mit Hilfe eines Datenverarbeitungssystems erstellt worden, hat die Finanzbehörde im Rahmen einer Außenprüfung das Recht, Einsicht in die gespeicherten Daten zu nehmen und das Datenverarbeitungssystem zur Prüfung dieser Unterlagen zu nutzen.“

unterrichtet, um geeignete Verbildlichungsansätze für die jeweiligen Fragestellungen zu wählen (s. den Abschnitt „Methodenwissen“ unter Pkt. 2.2.). Die SRP-Vorlagen unterstützen diesen Prozess, indem sie dem Prüfungsduktus „vom groben Überblick zu den Details“ folgen, Anwendungsregeln der verwendeten Methoden beachten, interaktive Optionen für das nachträgliche Einpflegen von neuen Erkenntnisse bzw. Einlassungen (s. den Abschnitt „Gesamtbildprüfungen“ unter Pkt. 2.3. und das Beispiel unter Pkt. 4.2.9.1.) zur Verfügung stellen und eine ausreichende Dokumentation erleichtern. Der zielführende Einsatz der SRP-Vorlagen zum Suchen von Prüffeldern sowie zum Aufdecken und Abklären von Implausibilitäten gem. § 158 AO wird intensiv geschult. Zusätzlich besteht eine bundesweite Austauschplattform der Finanzverwaltung, über welche weiterführende Ideen zur Umsetzung der SRP-Prüfprinzipien auf spezifische Prüfungsthemen in Echtzeit geteilt werden können.

Damit ein reibungsloser Ablauf des dienststellenübergreifenden Prüfverfahrens nicht behindert wird, gibt es angepasste Fortbildungsformate für alle involvierten Bereiche der Finanzverwaltung wie z.B. die Rechtsbehelfs- sowie Bußgeld- und Strafsachenstellen.

Ziel der SRP-Einführung ist es nicht, andere Herangehensweisen und Methoden der Prüfungsdienste zu verdrängen. Vielmehr ergänzt ein systematisch mehrperspektivisches Gesamtbild über den Inhalt von Besteuerungsdaten zu Beginn einer Außenprüfung das Ergebnis einer Formalprüfung i.S. des § 158 AO.⁹⁹ Danach ermöglicht die SRP als schnelle und leistungsfähige Orientierungsmaßnahme eine sachgerechte Einschätzung zur angemessenen Prüfungstiefe und zum Einsatz arbeitsaufwendigerer Prüfungsmethoden wie z.B. dem Datenzugriff auf vorgelagerte Systeme „**DavoS**“ (niedersächsisches Prüfungsprojekt) oder einer statistischen Stichprobenanalyse mit Monetary Unit Sampling (MUS). I.R. von Plausibilitätsprüfungen nach § 158 AO können ebenso zeitintensivere Verprobungsverfahren wie Geldflussrechnungen oder Kalkulationen bedarfsgerecht eingesetzt werden, um den bisherigen Gesamteindruck im Verhältnis zum gesetzlich geforderten Gewissheitsmaß „Beanstandungsanlass“ (s. Pkt. 4.6.) ausreichend abzusichern.

Zur **Wahrung der Transparenz** im Umgang der Außenprüfungsdienste mit den neuen digitalen Möglichkeiten¹⁰⁰ sind zur SRP von Beginn an kontinuierlich Beiträge veröffentlicht worden.¹⁰¹ Indem mit IDEA® und den auf MS Excel® bzw. OpenOffice® basierenden SRP-

⁹⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist.“

¹⁰⁰ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 49: „Anspruch der Steuerpflichtigen auf einen effektiven Rechtsschutz [...] (Art. 19 Abs. 4, Art. 20 Abs. 3 GG).“

¹⁰¹ Erste Beschreibungen zur „Summarischen Risikoprüfung (SRP)“ finden sich in: *Wähnert*, „Summarische Risikoprüfung – Echtes Risikomanagement in der Finanzverwaltung mit Hilfe einer neuen Prüfungsmethode“, StBp 2008, 296. Aktuellere Darstellungen liefern neben dem Wikipedia-Eintrag unter https://de.wikipedia.org/wiki/Summarische_Risikoprüfung u.a.: *Becker*, „Außenprüfung digital – Prüfungsme-

Vorlagen ausschließlich frei zugängliche Programme verwendet werden, ist zudem die Reproduzierbarkeit der Prüfungsergebnisse gewährleistet. Zunehmende Datenmengen in Außenprüfungen können hingegen nicht der Finanzverwaltung als Problem nachlassender Klarheit angelastet werden¹⁰², weil dies eine zwangsläufige Folge der Umstellung auf die elektronische Verwaltung in der Wirtschaft ist (s. [Pkt. 2.1.](#)). Im Gegenteil fördert der offene Umgang der Finanzverwaltung mit ihren neuen Prüfungsansätzen medial wahrnehmbar die Einsicht in Wirtschaft und Beratung deren Aufgabe „Tax Compliance“ unter den hinzugewonnenen digitalen Optionen neu auszugestalten¹⁰³, wodurch der – im Abschnitt „Risikomanagement“ unter diesem Pkt. benannte – Hauptzweck der Besteuerungskontrolle „Prävention“ grundlegend gestärkt wird.

3.1. Die SRP als Orientierungshilfe zu Prüfungsbeginn

Haupteinsatzbereich der SRP ist die Erlangung eines inhaltlichen Überblicks über Besteuerungsdaten zu Beginn einer Außenprüfung mit dem **Erkennen und Eingrenzen betriebspezifischer Prüffelder**. Im Anschluss an die Formalprüfung, ob Besteuerungsunterlagen prüfbar sind, ist dieser Überblick die Voraussetzung für eine zielführende, effiziente Ressourcenverwendung.

Durch die systematisch mehrperspektivische Datenverbildlichung kann in wenigen Stunden ein aussagekräftiger Eindruck zu Vertiefungsnotwendigkeiten auch bei sehr großen Datenmengen erarbeitet werden. Regel- und Musterabweichungen in den Entwicklungen, Abhängigkeiten und Häufigkeitsverteilungen der untersuchten Unternehmensdaten legen die aus Risikogesichtspunkten sinnvollsten Intensivierungsbereiche offen. Dabei sind die Funktionalitäten der SRP-Vorlagen so ausgestaltet, dass aufgefundene Regel- und Musterstörungen unter wenig Aufwand bis zu den Verursachern nachverfolgt werden können.

Das Effizienzgebot aus § 7 BpO, „[d]ie Außenprüfung [...] auf das Wesentliche abzustellen [und i]hre Dauer [...] auf das notwendige Maß zu beschränken“, erfordert ein **systemati-**

thoden im Fokus (Teil II)“, DStR 2016, 1430, *Nowotzin/Teutemacher*, „Neue Wunderwaffe der BP: Summarische Risikoprüfung – Die neuen Prüfungstechniken und wie Steuerberater diese besser verstehen“, IWW BBP Sonderheft 2017, *Harle/Olles*, „Die moderne Betriebsprüfung“, NWB Verlag, ab 3. Aufl. 2017, 221, NWB Betriebsprüfungs-Kartei, „Summarische Risikoprüfung (SRP)“, Stand: Februar 2017; Oberfinanzdirektion Nordrhein-Westfalen, in Zusammenarbeit mit der Finanzverwaltung Schleswig-Holstein, „Summarische Risikoprüfung (SRP) und Neue Prüfungstechnik (NiPt) – Wichtige Begriffe“, juris-Dokumenten-ID „bpnw-FMBP00000777“ (Pfad: „Fachinfosystem Bp > Allgemeine Prüfungshinweise > Arbeitspapiere und Leitfäden“), downloadbar auch unter: <https://taxtech.blog/2018/03/09/glossar-zur-summarischen-risikopruefung-srp-und-neuen-pruefungstechnik-nipt>.

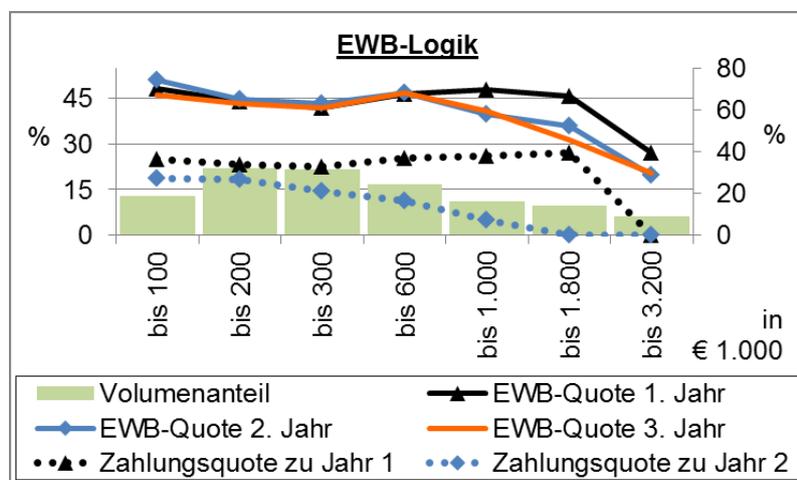
¹⁰² So einseitig geschehen in: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 50.

¹⁰³ Vgl. z.B.: *Liekenbrock*, „Digital Tax Compliance – Praxishinweise zur technischen Umsetzung von Tax Compliance Management Systemen“, Ubg 2018, 43, sowie *Ehlers*, „Tax Compliance – Neue Beratungsfelder bei Betreuung kleiner und mittlerer Unternehmen“, NWB 18/2012, 1535.

sches Vorgehen und den Einsatz leistungsfähiger Prüfwerkzeuge. Unter den Bedingungen der heutzutage üblichen elektronischen Betriebsverwaltung mit rasant ansteigenden Datenmengen und zusätzlichen spezifischen Fehler- und Manipulationsrisiken (s. [Pkt. 2.1.](#)) sind digitale Hilfsmittel in der Außenprüfung unverzichtbar. Der Methodenreichtum der digitalen Datenanalyse birgt jedoch die Gefahr, dass (auch) ungeeignete Verfahren eingesetzt bzw. unsystematische Abläufe gewählt werden.

Die SRP stellt deshalb als Basis ein ausreichendes Methodenwissen zu den Einzelmethoden mit deren Anwendungsmöglichkeiten und Eignungen zur Verfügung (s. [Pkt. 2.2.](#)). Davon ausgehend wird der planvolle Prüfungsduktus „vom groben Überblick zu den Details“ vorgegeben, der zum einen das Aufspüren von Systemfehlern sowie Einzelprüffeldern ermöglicht und zum anderen die zügige Eingrenzung und Abklärung unterstützt. Mit dem Prüfungsduktus und den daran angelehnten SRP-Vorlagen wird das Prüfungspersonal in die Lage versetzt, seine Erfahrung und Intuition zielführend in den komplexer gewordenen Prüfungssituationen zu nutzen.

Die fortgeschrittene freiere Anwendung der SRP-Prüfungsprinzipien in Form von individuellen Lösungen für **komplizierte Einzelprüfungsfragen** vermittelt in der Visualisierungsbildung den zielführenden Einsatz von IDEA® und Tabellenkalkulationsprogrammen



(s. [Pkt. 5.](#)) So bewirkt eine fokussierte Verbildlichung unübersichtlicher Verhältnisse wie im abgebildeten Beispiel zur fortlaufenden Logik der Einzelwertberichtigungen (EWB) einer Bank i.d.R. eine erheblich klarere rechtliche Entscheidungsgrundlage.¹⁰⁴

In dem auf die wesentlichen

Informationen konzentrierten Diagramm ist gut zu erkennen, dass der bis auf die höchsten Kredite recht einheitliche EWB-Anteil im ersten Jahr durch die ebenso gleichmäßige Zahlungsquote im Folgejahr bestätigt wird. Demgegenüber wurde die **EWB-Quote der drei höchsten Kreditklassen im zweiten Jahr deutlich gesenkt**, woraufhin die **nachträgliche Begleichung in diesen Bereichen auf (nahezu) Null** abfällt. Daraus sind **im dritten Jahr offensichtlich keine Konsequenzen gezogen** worden, denn der EWB-Anteil wurde für die hohen Kreditklassen weiterhin niedriger gehalten. Es lässt sich demnach im Überblick eindeutig

¹⁰⁴ Interessanterweise werden zur Darstellung von Beziehungen verbundener Unternehmen üblicherweise Organigramme als Verbildlichungen verwendet, wohingegen die eigentliche Prüfung der viel komplexeren Unternehmensdaten auf der Basis von Tabellen geschieht.

ablesen, dass die Bewertung der Kredite unterschiedlicher Größenklassen nicht nach denselben Maßstäben vorgenommen wurde. Nach den ständig zunehmenden Erfahrungen der Prüfungspraxis auch in dem Bereich komplexer Einzelprüffelder der Groß- und Konzernbetriebsprüfung treten durch die gezielte Verbildlichung häufig bisher ungeahnte Prüfungsfragen hervor.

3.2. Die SRP im besonderen Kontext des § 158 AO

Gesamtbildprüfungen mit der SRP – insbesondere zur Vollständigkeit von (Bar-)Einnahmen – stellen zwar nur einen kleinen Ausschnitt der Anwendungsbreite dar, werden von KritikerInnen aber gerne in den Fokus gerückt¹⁰⁵, denn die Aberkennung der Beweiskraft von Besteuerungsaufzeichnungen nach § 158 AO löst gem. § 162 Abs. 2 S. 2 AO die Amtspflicht der Schätzung aus.

Weil die Besteuerung einen schweren Eingriff in die Eigentumsrechte der BürgerInnen bedeutet, soll das **Verifikationsprinzip** die Gleichmäßigkeit in der Anwendung der Steuergesetze gewährleisten.¹⁰⁶ Dazu führt das ordnungswidrige Aufzeichnen und Aufbewahren von Besteuerungsunterlagen, falls es die Prüfung durch „eine[n] sachverständigen Dritten innerhalb angemessener Zeit“¹⁰⁷ verhindert, genauso zur Schätzung der Besteuerungsgrundlagen wie ein „Anlass [...], ihre sachliche Richtigkeit zu beanstanden“¹⁰⁸. Durch dieses außergewöhnliche Beweislastverhältnis soll sichergestellt werden, „dass der ‚Beweisverderber‘ oder ‚Beweisvereiteler‘ aus seinem Verhalten keinen Vorteil ziehen darf [...] Zur Vermeidung [...] sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt.“¹⁰⁹

Besonders komplexe Gesamtbilder wie bspw. zur Vollständigkeit von (Bar-)Einnahmen lassen sich aus den unvollständigen und nicht objektiven Dokumentationen sowie Erinnerungen der Vergangenheit nur unter Unsicherheiten rekonstruieren. Deshalb sieht der Gesetzgeber die „Beweiskraft der Buchführung (und Aufzeichnungen)“ nach § 158 AO bereits dann als widerlegt an, wenn es „nach Verprobung usw. unwahrscheinlich ist, dass das aus-

¹⁰⁵ Vgl. z.B.: IWW online, PStR Praxis Steuerstrafrecht v. 17.9.2015: „Summarische Risikoprüfung: Das neue Einfallstor für Steuerstrafverfahren?“, nachzulesen unter:
<https://www.iww.de/pstr/schwerpunktthema/betriebspruefung-summarische-risikopruefung-das-neue-einfallstor-fuer-steuerstrafverfahren-f88749>.

¹⁰⁶ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

¹⁰⁷ § 145 Abs. 1 AO.

¹⁰⁸ § 158 AO.

¹⁰⁹ BFH v. 15.2.1989, BStBl. II 1989, 462.

gewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt¹¹⁰. Aus der Konkretisierung des (juristischen) Strengbeweises gem. BGH¹¹¹ in Verbindung mit den probabilistischen Theorien zur Vereinheitlichung von Aussagen unter Unsicherheit (Wahrscheinlichkeitsaussagen) der Gutachtenpraxis¹¹² lässt sich das gesetzliche Gewissheitsmaß des „**Beanstandungsanlasses**“ bei ca. 95 % einordnen (s. [Pkt. 4.6.](#)).

Wie die herkömmlichen Verprobungen „Vermögenszuwachs-“, „Geldverkehrsrechnung“ und „Kalkulation“ sind Gesamtbildergebnisse einer SRP an diesem Maßstab des § 158 AO zu messen. Die Vorteile systematisch mehrperspektivischer Prüfungsnetze liegen neben der höheren Aufdeckungsquote vor allem darin, dass entdeckte Unschlüssigkeiten über ergänzende Datenverbildlichungen abgesichert werden. Zusätzlich lassen sich neue Erkenntnisse z.B. durch Einlassungen interaktiv (s.a. den Abschnitt „Gesamtbildprüfungen“ unter [Pkt. 2.3.](#) und [Pkt. 4.2.9.](#)) einpflegen. Das resultierende Gesamtbild ist somit stets auf dem aktuellen Prüfungsstand und besteht nach umfangreichen Erfahrungen mit der SRP¹¹³ bei beanstandeter Plausibilität i.d.R. aus mehreren Indizien (s.a. das Beispiel unter [Pkt. 4.4.](#)). Damit ist die Wahrscheinlichkeit, dass die Schwelle des § 158 AO erreicht wird, deutlich größer, wie es die BGH-Rechtsprechung allgemein mit der Bedeutung von Indiziengesamtheiten festgestellt hat¹¹⁴.

Sofern die Schätzungsnotwendigkeit gem. § 162 Abs. 2 S. 2 AO gegeben ist, bieten die SRP-Vorlagen zwei **Schätzungsoptionen** an, um die bereits gewonnenen tiefreichenden Erkenntnisse zu den Besteuerungsdaten für die Ermittlung einer sachgerechten Schätzungshöhe zu nutzen:

¹¹⁰ Anwendungserlass zu § 158 AO.

¹¹¹ Urt. v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 26.

¹¹² Köller/Nissen/Rieß/Sadorf, „Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004.

¹¹³ Aus dem (Flächen-)Einsatz der SRP seit 2008 in inzwischen drei Viertel der deutschen Bundesländer durch zurzeit mehr als 2.000 PrüferInnen sind viele tausend Fälle hervorgegangen, die infolge der indiziellen Beurteilung der inhaltlichen Plausibilität zur Schätzung geführt haben. Dabei wurde die Einschätzung in Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist“ nicht nur regelmäßig dahingehend bestätigt, dass formelle und materielle Mängel – sofern letztere noch feststellbar sind – gemeinsam auftreten, sondern auch insofern, dass Indizien üblicherweise nicht einzeln vorkommen.

¹¹⁴ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweismwürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

- © Die **Quantilsschätzung** ist eine Vollschätzungsmethode, um auf Basis der regelbasierten Zeitreihenanalysen intern die objektivierte Leistungsfähigkeit abzuleiten und folgt damit dem Gebot des BFH, „in dem gegebenen Schätzungsrahmen an die oberste Grenze zu gehen“¹¹⁵. Dabei gründet sich die Quantilsschätzung nicht auf einer allgemeingültigen Gesetzmäßigkeit, weil Steuerhinterziehungen nicht einheitlich verlaufen, sondern nutzt für die notwendigen Schritte, um aus Zeitreihenanalysen ein Schätzungsergebnis zu gewinnen, die gängigen, anerkannten Methoden aus Statistik und BWL (s. [Pkt. 4.8.](#)).
- © Mit der **Verteilungsanalyse** als Häufigkeitsauszählung von täglichen Bareinnahmen nach Größenklassen eröffnet sich *in geeigneten Fällen* eine weitere Möglichkeit, intern eine Schätzungshöhe zu ermitteln. Da Wirtschaftsdaten nach den Erkenntnissen der Wissenschaft i.d.R. der Logarithmischen Normalverteilung (s. [Pkt. 4.3.4.1.](#)) folgen, kann aus dem Vergleich der beobachteten Werte mit der fallspezifisch rekonstruierten Logarithmischen Normalverteilung zumindest der Umfang eines – manipulationspsychologisch begründeten – stärkeren Eingriffs in höhere Werte abgeleitet werden. Damit ermittelt diese Form der Schätzung allerdings eher einen Mindestschaden, der dann den Anforderungen einer strafrechtlichen Schätzung entspräche (s. [Pkt. 4.9.](#)).

Tlw. wird inzwischen auch in der Literatur die größtmögliche Realitätsnähe von Gesamtbildanalysen mittels SRP betont.¹¹⁶ Nach derzeitigem Stand der Rechtsprechung werden vom X. Senat des BFH mit der „Drei-Stufen-Theorie“ die herkömmlichen Verprobungen „Vermögenszuwachs-“, „Geldverkehrsrechnung“ und „Kalkulation“ vorgezogen¹¹⁷, wobei die aufgeführten Kritikpunkte an der Quantilsschätzung ausnahmslos unzutreffend sind (s. *den Abschnitt „Finanzrechtsprechung und Kritikpunkte“ unter [Pkt. 4.8.](#)*). Dagegen haben einige Finanzgerichte sie als differenzierte griffweise Schätzungsmethode anerkannt.¹¹⁸

¹¹⁵ Urt. v. 9.3.1967, BStBl. III 1967, 349; aktuell bestätigt von: *Nöcker (RiBFH)*, „Zeitreihenvergleich im Gleitschlitzen versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3549.

¹¹⁶ *Pump/Heidl*, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenbuchführung“, StBp 2014, 162, 163: „Dabei werden durch umfassende Erfassung und Darstellung der betrieblichen Verhältnisse [...] die Besteuerungsgrundlagen extrem realitätsnah ermittelt.“

¹¹⁷ Urt. v. 25.3.2015, BStBl. II 2015, 743, 4. Leitsatz.

¹¹⁸ Vgl. z.B.: FG Hamburg v. 26.8.2016 – 6 V 81/16, v. 31.10.2016 – 2 V 202/16 und v. 18.7.2017 – 6 V 119/17 sowie FG Berlin-Brandenburg v. 9.1.2017 – 4 V 4265/15 u. v. 30.6.2017 – 10 V 10219/16: „Soweit das FG Berlin-Brandenburg [...] demgegenüber die Quantilsschätzung verworfen und eine pauschale Erhöhung der Umsätze der dortigen Antragstellerin um 10 % für vorzugswürdig gehalten hat, folgt der Senat dem nicht, weil die Quantilsschätzung eine deutlich differenziertere Schätzungsmethode darstellt.“

4. Die Verwendung der SRP-Vorlagen (Basis-SRP)

Hauptzweck der SRP ist die Vermittlung eines umfassenden Verständnisses visualisierender Prüfungsansätze, um dadurch zu Beginn einer Außenprüfung zügig eine gute Orientierung in umfangreichen Besteuerungsdaten zu erlangen und geeignete Methoden für die zielführende Bearbeitung unterschiedlichster Fragestellungen im Prüfungsalltag einzusetzen (s. [Pkt. 2.](#)). Hierfür stellen die SRP-Vorlagen einen abgestimmten Werkzeugkasten zur Verfügung, um die Vorteile systematisch mehrperspektivischer Prüfungsnetze auch bei geringen analytischen Kenntnissen erfahren zu können.

Die unter diesem Pkt. 4. beschriebene Basisanwendung mit Hilfe der SRP-Vorlagen erfolgt exemplarisch an dem Anwendungsbereich „Prüfung von (Bar-)Einnahmen auf Vollständigkeit“, weil hier der inhaltliche Zusammenhang der vielen regel- und musterbasierten Datenverbildlichungen gem. SRP-Prüfungsduktus gut nachvollzogen werden kann. Für die allgemeine Anwendung der SRP werden unter [Pkt. 5.](#) einige Beispiele vorgestellt und besteht eine eigene Visualisierungsfortbildung mit Fallbeispielen zu diversen Prüffeldern.

Beiden SRP-Vorlagen ist als erstes Tabellenblatt eine **Kurzanleitung** vorgeordnet, welche jeweils den Versionsstand, im weißen Bereich eine kurze Beschreibung der Vorlage und generelle Bedienhinweise sowie im grünen bis roten Teil eine knappe Erläuterung der einzelnen Arbeitsblätter bzw. Analyseschritte entlang des SRP-Prüfungsduktus enthält. Auf den einzelnen Arbeitsblättern haben die ausfüllbaren Felder einen weißen Hintergrund (*Die anderen Felder sind mit einem Passwort überschreibgeschützt.*) und werden die Funktionalitäten über Ankreuz- oder Auswahlelemente gesteuert. Anhand der roten Ecken oben rechts in den betreffenden Feldern sind Kommentare zu erkennen, die entweder in blauer Schrift als Erklärungs- oder in roter als Warnhinweis helfen sollen. Viele Rechenfelder beinhalten Plausibilitätsprüfungen, ob die eingegebenen Daten sinnvoll erscheinen und wichtige Arbeitsschritte beachtet wurden. Beim Einsatz von Wahrscheinlichkeitstests werden deren Anwendungsregeln beachtet bzw. wird auf die Grenzen der Auswertbarkeit hingewiesen. Trotzdem können diese Hilfsmittel der SRP-Vorlagen nicht das ausführliche SRP-Handbuch ersetzen.

Sämtliche Arbeitsblätter beider SRP-Vorlagen sind **ausdruckvorbereitet** und geben dabei alle zur Reproduktion erforderlichen Informationen, die Dateibezeichnung, das Ausdrucksdatum sowie etwaige Datenbeeinflussungen (bspw. Einzeldatenkorrektur in der SRP-Vorlage BWA oder Einschränkung der untersuchten Daten in der SRP-Vorlage Stochastik) aus. Somit benötigen Stpfl. oder BeraterInnen zur Überprüfung der Ergebnisse lediglich die

verwendeten Daten z.B. als Export von IDEA® in eine leere MS Excel®- oder OpenOffice®-Mappe. Zudem lassen sich auf jedem Auswertungsblatt in dem jeweiligen „Bemerkung(en)“-Feld freie Eintragungen ergänzen. Die Herausgabe der Vorlagendateien ist somit zur Erfüllung des „Anspruchs der Steuerpflichtigen auf einen effektiven Rechtsschutz durch die Gerichte (Art. 19 Abs. 4, Art. 20 Abs. 3 GG)“¹¹⁹ nicht erforderlich!

Neue Vorlagenversionen werden mit einer kurzen Übersicht über die Aktualisierungen ausgegeben, wodurch eine schnellere Orientierung ermöglicht werden soll.

Die **Nutzung der SRP-Vorlagen** beginnt damit, die zu untersuchenden Besteuerungsdaten mit ihren Spalten (Merkmalen) „Datum“ und „Betrag“ aus IDEA® oder z.B. einem Tabellenkalkulationsprogramm in die dafür vorgesehenen Felder der Arbeitsblätter „Vergleichsdaten“ (*SRP-Vorlage BWA*) bzw. „Daten“ (*SRP-Vorlage Stochastik*) hineinzukopieren:

- ☰ zwei Datenreihen, deren Entwicklungen und Abhängigkeit untersucht werden sollen, in die SRP-Vorlage BWA wie z.B. Wareneinkauf und -umsatz, operative Kosten(stellen) und eine Erfolgsgröße, Vertrags Soll und -ist oder die Datenbestände vor und nach einer Datenschnittstelle bzw.
- ☰ eine Datenreihe, deren Häufigkeitsverteilungen auf Muster analysiert werden sollen, in die SRP-Vorlage Stochastik wie z.B. eine Vielzahl an Kostenpositionen zum zügigen Auffinden von betriebsspezifischen Prüffeldern oder Tagesbareinnahmen für die Prüfung auf Manipulationsspuren.

Vorsicht: Beide SRP-Vorlagen haben nur ein begrenztes Fassungsvermögen der Eingabebereiche (s. den jeweiligen Haupthinweis auf den Tabellenblättern „Vergleichsdaten“ bzw. „Daten“). Sollen größere Datenbestände untersucht werden, gibt es die Möglichkeit über das Arbeitsblatt „eigene Notizen“ mit einer zwischengeschalteten Pivot-Tabelle zu arbeiten (s. den nächsten Abschnitt „fortgeschrittene Anwendung der SRP-Vorlagen“ zu diesem Pkt.).

Zur **Kontrolle der korrekten Datenübernahme** wird in der SRP-Vorlage BWA nach dem Einfügen die **Datensatzzahl der jeweiligen Datenreihe** angegeben. Auch die jährlichen Betrags-

Datum	Betrag
853	853
02.01.2002	105,16

summen auf dem Tabellenblatt „Jahreszahlen“ gewährleisten eine schnelle Überprüfung, dass die richtigen Daten vollständig in die Vorlage hineinkopiert worden sind. Bei der SRP-Vorlage Stochastik werden auf dem Tabellenblatt „Daten“ im Bereich „statistische Kenngrößen“ die erforderlichen Parameter wie „Anzahl“, „Summe“ etc. für eine einfache Gegenkontrolle angezeigt.

I.R. von **Plausibilitätsprüfungen** mit der SRP-Vorlage BWA auf die kaufmännische Handlungslogik von Betriebsdaten wie der Abhängigkeit des Wareneinkaufs von den Waren-

¹¹⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 49.

erlösen werden üblicherweise – entsprechend der betriebswirtschaftlichen Sicht – **Nettozahlen** verwendet. Je leistungsnäher dabei die Datumswerte z.B. infolge der Verwendung eines elektronischen Warenwirtschaftssystems sind, desto aussagekräftiger sind die Schlüssigkeitsauswertungen.

Zur speziellen Untersuchung von (Bar-)Einnahmen auf **Manipulationsspuren** müssen diejenigen Daten in die SRP-Vorlage Stochastik eingefügt werden, die – nach dem jeweiligen Kenntnisstand zu den Aufzeichnungsprozessen – Ausgangspunkt menschlicher Eingriffe gewesen sein können: i.d.R. Bareinnahmen der konsolidierten Kasse(n) als **täglich aufsummierte Bruttowerte**¹²⁰.

Weitere Eingaben in die SRP-Vorlagen sind:

- Die Eingaben zu „Steuerpflichte/r“ und „Steuernummer“ werden automatisch auf alle Auswertungsblätter übernommen.
- In der SRP-Vorlage BWA lassen sich die Überschriften der beiden vorbenannten Datenreihen „Einkauf“ und „Umsatz“ auf dem Tabellenblatt „Vergleichsdaten“ sowie des Verhältnisses „RGAS“ auf dem Tabellenblatt „Quartalsrelationen“ überschreiben und werden in alle Auswertungen übertragen.
- Beginn und Ende des Prüfungszeitraums müssen auf dem Tabellenblatt „Vergleichsdaten“ der SRP-Vorlage BWA eingegeben werden, wodurch die Auswertungen auch unproblematisch auf Teilzeiträume beschränkt werden können. Beide Vorlagen unterstützen nach entsprechender Eingabe auf dem Tabellenblatt „Vergleichsdaten“ (*SRP-Vorlage BWA*) bzw. „Daten“ (*SRP-Vorlage Stochastik*) das abweichende Wirtschaftsjahr.
- Neben den Eingabespalten „Datum“ und „Betrag“ für die beiden Vergleichsdatenreihen verfügt die SRP-Vorlage BWA jeweils über eine Korrektur- und eine Bemerkungsspalte, um punktuelle Berichtigungen von zeitlichen Zuordnungen oder Wertinhalten einzelner Datensätze (Sachverhalte) interaktiv vorzunehmen (s. [Pkt. 4.2.9.](#)).
- Mit dem freien Feld oberhalb der Spalte „Datum“ zur ersten Vergleichsdatenreihe kann diese insgesamt um bis zu 20 Tage vor- oder rückdatiert werden, um die Auswirkungen auf die Analyseergebnisse interaktiv zu simulieren (s. [Pkt. 4.2.9.](#)).
- Auf dem Tabellenblatt „Jahreszahlen“ der SRP-Vorlage BWA können i.R. von Plausibilitätsprüfungen des Warenflusses jährliche Bestandsveränderungen und Entnahmen sowie Richtsatzwerte eingetragen werden. Die Entnahmewerte lassen sich auf dem Tabellenblatt „Monatszeitreihen“ über ein Ankreuzfeld auf die (vorhandenen) Quartale und Monate gleichmäßig aufteilen. Eine entsprechende Funktion ist für die Bestandsverän-

¹²⁰ Sollen aufgrund eines konkreteren Verdachts einzelne Kassen oder die Einnahmen nach Umsatzsteuersatz getrennt untersucht werden, wird jeweils eine eigene Vorlage verwendet.

derungen nicht vorgesehen, weil eine gleichmäßige Verteilung nicht den üblichen Verhältnissen entspricht. Über die interaktiven Optionen (s. [Pkt. 4.2.9.](#)) können jedoch punktuelle Umverteilungen vorgenommen werden.

- ☒ Für die retrograde Rekonstruktion des Warenbestands kann auf dem Tabellenblatt „Bestandsentwicklung“ ein (geschätzter) Anfangsbestand eingetragen werden.

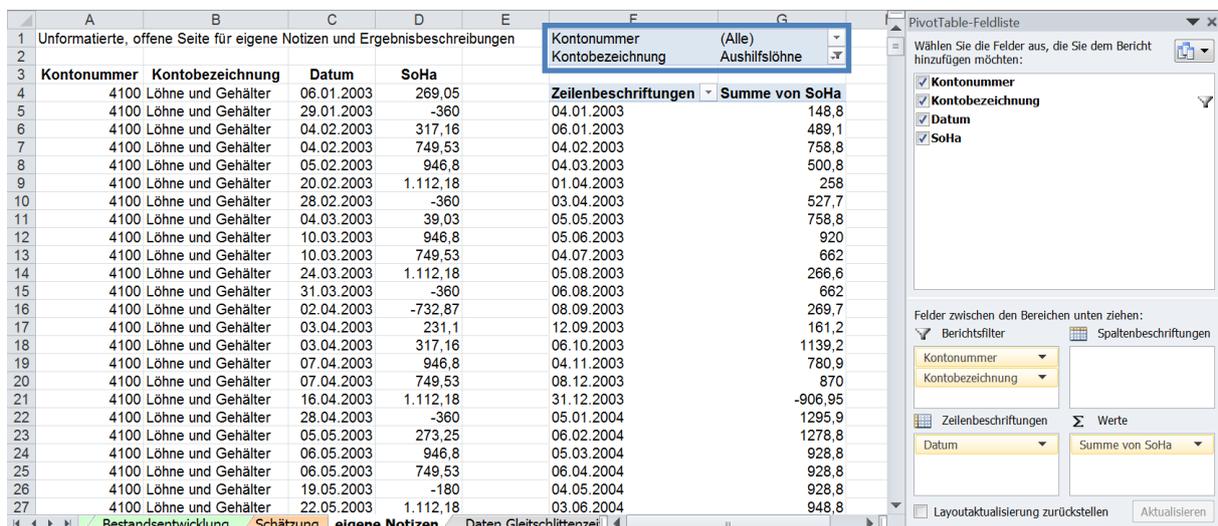
Die **SRP-Vorlage Stochastik** bietet zur gezielten Vertiefung von Musterstörungen als risikoorientierte Prüffelder sowie zur Absicherung entdeckter Manipulationsmuster eine **zweistufige Eingrenzung der untersuchten Daten** an: Die Einschränkungen auf Jahre, einen Teilzeitraum und/oder Wochentage des Tabellenblatts „Daten“ wirken grundsätzlich auf alle Arbeitsblätter, während die jeweiligen Beschränkungen einzelner Analysen wie z.B. auf dem Tabellenblatt „digZA“ nur dort gelten. Einzelne Datensätze lassen sich zudem nachvollziehbar auf dem Tabellenblatt „Daten“ über die Ankreuzfelder links neben den importierten Daten ausschließen. Alle grundsätzlichen Einschränkungen des Tabellenblatts „Daten“ werden auf den einzelnen Auswertungsblättern angezeigt und dienen i.R. der Manipulationsanalyse einer verbesserten und abgesicherten Aufdeckung¹²¹.

Der **Prüfungsduktus** der vielen regel- und musterbasierten Einzelmethoden folgt in der SRP-Vorlage BWA dem durch die Reiterabfolge vorgegebenen Ablauf, um „vom groben Überblick zu den Details“ sowohl grobe Systemauffälligkeiten als auch punktuelle Prüffelder zügig und zielgenau aufzufinden bzw. die Plausibilität i.S. des § 158 AO erst nach Ausschluss (möglicher) Störeinflüsse zu bewerten. Weil eine derart stringente Reihenfolge der musterbasierten Analysen i.R. der Manipulationsprüfung nicht möglich ist, enthält die SRP-Vorlage Stochastik auf dem Tabellenblatt „Kurzanleitung“ ein Prüfungsschema, welches sich auf die Aufdeckung zweier praxisrelevanter Manipulationsszenarien („erfundene Einnahmewerte“ und „stärkere Verkürzung hoher Einnahmen“) konzentriert. Dadurch soll als Ergebnisabsicherung ein weitreichender Ausschluss betrieblicher Sonderursachen gewährleistet werden.

Der beweisverstärkende Vorteil von Indiziengesamtheiten im Zusammenhang mit Plausibilitätsprüfungen gem. § 158 AO durch den Einsatz systematisch mehrperspektivischer Prüfungsnetze wie SRP setzt die **zusammenführende Dokumentation** aller Einzelergebnisse von der Formalprüfung über die einzelnen Auswertungen bis zum interaktiven Umgang mit Einlassungen voraus.

¹²¹ Der schnelle Wandel im Bereich der elektronischen Betriebsverwaltung lässt bei Manipulationen Veränderungen im Vorgehen (z.B. Umstieg von manueller Verkürzung zu Zappern) erwarten, was eine zeitliche Beschränkung möglicher Manipulationsspuren nach sich zieht. Außerdem können die Dateneingrenzungen eine Differenzierung zwischen Manipulation und Unterschlagung ermöglichen.

Als **fortgeschrittene Anwendung der SRP-Vorlagen** lassen sich einerseits deutlich größere Datenbestände als im jeweilig vordefinierten Datenbereich (*Tabellenblätter „Vergleichsdaten“ in der SRP-Vorlage BWA und „Daten“ in der SRP-Vorlage Stochastik“*) untersuchen und andererseits Unterscheidungskriterien einbinden, indem mit Pivot-Tabellen auf dem Tabellenblatt „eigene Notizen“ gearbeitet wird: Dazu werden die zu untersuchenden Daten inkl. zusätzlicher Differenzierungsmerkmale (in der folgenden Grafik: „Kontonummer“ und „-bezeichnung“) im Arbeitsblatt „eigene Notizen“ abgelegt. Anschließend werden die Einzeldaten durch Einfügen einer Pivot-Tabelle geordnet. Dabei sorgt die Zuordnung des Datums als „Zeilenbeschriftung“ (per „drag and drop“) für eine tägliche Aufsummierung, so dass eine begrenzte Anzahl an Tageswerten resultiert. Nach Zuweisung der Wertinformation (hier: „SoHa“) werden die zusätzlichen **Merkmale „Kontonummer“ und „Kontobezeichnung“ als Berichtsfiler** (bzw. Datenschnitte¹²²) gewählt:



Kontonummer	Kontobezeichnung	Datum	SoHa	Summe von SoHa
4100	Löhne und Gehälter	06.01.2003	269,05	148,8
4100	Löhne und Gehälter	29.01.2003	-360	489,1
4100	Löhne und Gehälter	04.02.2003	317,16	758,8
4100	Löhne und Gehälter	04.02.2003	749,53	500,8
4100	Löhne und Gehälter	05.02.2003	946,8	258
4100	Löhne und Gehälter	20.02.2003	1.112,18	527,7
4100	Löhne und Gehälter	28.02.2003	-360	758,8
4100	Löhne und Gehälter	04.03.2003	39,03	920
4100	Löhne und Gehälter	10.03.2003	946,8	662
4100	Löhne und Gehälter	10.03.2003	749,53	266,6
4100	Löhne und Gehälter	24.03.2003	1.112,18	662
4100	Löhne und Gehälter	31.03.2003	-360	269,7
4100	Löhne und Gehälter	02.04.2003	-732,87	161,2
4100	Löhne und Gehälter	03.04.2003	231,1	1139,2
4100	Löhne und Gehälter	03.04.2003	317,16	780,9
4100	Löhne und Gehälter	07.04.2003	946,8	870
4100	Löhne und Gehälter	07.04.2003	749,53	-906,95
4100	Löhne und Gehälter	16.04.2003	1.112,18	1295,9
4100	Löhne und Gehälter	28.04.2003	-360	1278,8
4100	Löhne und Gehälter	05.05.2003	273,25	928,8
4100	Löhne und Gehälter	06.05.2003	946,8	928,8
4100	Löhne und Gehälter	06.05.2003	749,53	928,8
4100	Löhne und Gehälter	19.05.2003	-180	928,8
4100	Löhne und Gehälter	22.05.2003	1.112,18	948,8

Nun können die Tagessummen der Pivot-Tabelle mittels Verweisformel¹²³ auf das jeweilige Datenblatt der SRP-Vorlagen verlinkt werden (**Achtung: Die Gesamtsumme der Spalte muss vorher über die Funktion „PivotTable-Optionen.../Summen & Filter/Gesamtsummen“ ausgeblendet werden**). Im Ergebnis kann die Zuweisung der zu untersuchenden Daten über die Berichtsfiler (bzw. Datenschnitte) ausgewählt und jederzeit umgestellt werden. Soll also im abgebildeten Beispiel der Zeitreihenvergleich der Aushilfslöhne zu den (direkt in das Ta-

¹²² Während die älteren Berichtsfiler voneinander unabhängig sind, erkennen Datenschnitte automatisch den Zusammenhang. So zeigt der Berichtsfiler „Kontonummer“ im Beispielsfall keine Einschränkung an, obwohl über die Kontobezeichnung nur die Aushilfslöhne ausgewählt wurden. Bei Datenschnitten würde die Auswahl unter „Kontobezeichnung“ auf die Anzeige zur Kontonummer durchwirken.

¹²³ In der SRP-Vorlage BWA reicht die einfache Verlinkung „=‘eigene Notizen‘!Feldadresse“, während in der SRP-Vorlage Stochastik eine Absicherung gegen leere Felder erfolgen muss: „=WENN('eigene Notizen'!Feldadresse="" ; "" ; 'eigene Notizen'!Feldadresse). Im Anschluss werden die ersten beiden Verlinkungen im vordefinierten Datenbereiche der SRP-Vorlagen bis zur maximalen Anzahl an Tagen (bei 3 Prüfungsjahren: max. 1.096) nach unten kopiert [s.a. die Arbeitshilfe „Erweiterte SRP Vorlagennutzung für zusätzliche Unterscheidungsmerkmale“ auf dem KTZ-Wiki-Server (Intranet) im Bereich „KTZWiki > PRUEFUNGS Web > Projekt_IDEA > SRP – Summarische Risikoprüfung“].

bellenblatt „Vergleichsdaten“ hineinkopierten) Umsätzen auf die Festangestellten ausgeweitet werden, muss lediglich die Einstellung des Berichtsfilters über die Auswahlfläche ergänzt werden. Danach sind alle BWA-Auswertungen umgehend an die neue Prüfung der Entwicklungen und Abhängigkeit angepasst.

4.1. Formalprüfung: Ordnungsmäßige (prüfbare) Aufzeichnungen

Als Anleitung für (systematisch visualisierende) Prüfungstechnik steht das umfangreiche Thema „(zeitgemäße) Formalprüfung“ – trotz seiner grundsätzlichen Bedeutung – nicht im Vordergrund des SRP-Handbuchs; dafür sei hier auf Werke mit umfassenden, detaillierten Ausführungen verwiesen, insbesondere:

- 📖 Achilles, „Kassenführung – Bargeschäfte sicher dokumentieren“, Verlag der DATEV e.G., 2018 (2. Auflage).
- 📖 Brinkmann, „Schätzungen im Steuerrecht: Fälle – Methoden – Vermeidung – Abwehr (Steuerrecht und Steuerberatung, Band 49)“, Erich Schmidt Verlag, 2017 (4. Auflage).
- 📖 Teutemacher, „Handbuch zur Kassenführung: Praxishandbuch für die rechtssichere Umsetzung“, NWB Verlag, 2018 (2. Auflage).
- 📖 Teutemacher, „Praxis-GoBD: Handbuch zur praktischen Umsetzung der Grundsätze zur ordnungsmäßigen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)“, NWB Verlag, geplant für Dez. 2018.

An dieser Stelle sollen lediglich einige grundsätzliche Erläuterungen zur **teleologischen**¹²⁴ bzw. **funktionalen**¹²⁵ **Auslegung** der abstrakten Ordnungsvorgaben erfolgen, denn durch die verbreitete Umstellung der Betriebsverwaltungen auf eine elektronische Arbeitsweise¹²⁶ haben sich die Steuerausfallrisiken – und damit die Ordnungsmäßig-



¹²⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 12.

¹²⁵ FG Hamburg v. 26.8.2016 – 6 V 81/16, 1. Leitsatz.

¹²⁶ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876 u. 877: „Die Buchführung hat sich von der Durchschreibebuchführung hin zu einer umfassenden Datenverarbeitung entwickelt, die elektronische Form ist nunmehr die übliche. [...] Diese Entwicklung geschah auf Betreiben und im Interesse der Unternehmen, weil es betriebswirtschaftlich notwendig war und ist, diese Einzeldaten zu erzeugen, z.B. für ein aussagekräftiges Controlling oder zur Arbeitsprozesssteuerung. [...] Dabei steigt die Anzahl elektronisch überwachter Prozesse stetig.“

keitsprüfung – fundamental verändert.

Ziel der Verknüpfung der „Beweiskraft der Buchführung (und Aufzeichnungen)“ nach § 158 AO mit strengen Ordnungsregeln für die steuerliche Aufzeichnung und Aufbewahrung gem. § 140 ff. AO ist es, das **Verifikationsprinzip**¹²⁷ zu stützen. Danach ist die Prüfbarkeit von Besteuerungsgrundlagen durch einen „sachverständigen Dritten innerhalb angemessener Zeit“¹²⁸ Grundvoraussetzung für das Vertrauen in die Steuererklärung, weil ein „Beweisverderber“ oder „Beweisvereitler“ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO).¹²⁹ Entsprechend müssen gravierende Formalmängel, die den Prüfungszugang entscheidend behindern, sanktioniert werden und führen gem. § 162 Abs. 2 S. 2 AO zur Amtspflicht der Schätzung.¹³⁰

Bereits mit der Reichsabgabenordnung¹³¹ wurden als **abstrakte Ordnungsvorgaben** die Kriterien „vollständig, richtig, zeitgerecht und geordnet“¹³² sowie „Gewissheit der Originalität“¹³³ definiert, welche die Nachvollziehbarkeit für einen „sachverständigen Dritten innerhalb angemessener Zeit“¹³⁴ sicherstellen sollen. Die Merkmale zielen dabei nicht nur auf das retrograde Auffinden eines einzelnen Vorgangs mit seinen Belegen ab, sondern gleichzeitig auf die Aussagefähigkeit des Gesamtbilds, weil es „weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann“¹³⁵. Die abstrakte Fassung des Ordnungsrahmens verfolgt die Wahlfreiheit der Aufzeichnungs- und Aufbewahrungsweise, was gerade im Hinblick auf die sich im-

¹²⁷ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

¹²⁸ § 145 Abs. 1 AO.

¹²⁹ BFH v. 15.2.1989, BStBl. II 1989, 462: „[...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

¹³⁰ Die „Drei-Stufen-Theorie“ aus dem BFH-Zeitreihenurteil [BStBl. II 2015, 743; ergänzend *Nöcker (RiBFH)*, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3553] widerspricht insofern mit der Abstufung von 2 („Buchführung formell nicht ordnungsgemäß, [...] aber materielle Unrichtigkeiten der Einnahmenerfassung nicht konkret nachgewiesen“, Rz. 63) auf 3 („Buchführung nicht nur formell, sondern auch materiell unrichtig“, Rz. 66) dem Verifikationsprinzip, weil gravierende Ordnungsmängel gerade das Nachweisen von materiellen Mängeln verhindern.

¹³¹ § 162 RAO.

¹³² § 146 Abs. 1 AO.

¹³³ § 146 Abs. 4 AO.

¹³⁴ § 145 Abs. 1 AO.

¹³⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5.

mer schneller wandelnde Technik realistisch ist. In der Konsequenz finden sich etliche Klarstellungen in der Rechtsprechung¹³⁶ oder in Erlassen¹³⁷.

Zusammenfassen lassen sich die Ordnungsvorgaben in der **abstrakten Kernregel** „übersichtliche, zeitgerechte und unveränderliche Einzelaufzeichnung“. Auch in der jüngsten BFH-Rechtsprechung ist hierbei verstärkt eine „funktionale“¹³⁸ Sichtweise zu erkennen.¹³⁹ Schon nach dem Gleichheitsgrundsatz in Art. 3 GG muss dieser Grundsatz auch für die Einnahmen-Überschuss-Rechnung gem. § 4 Abs. 3 EStG gelten¹⁴⁰, damit das Prinzip der Verifizierbarkeit gewahrt bleibt¹⁴¹. Dem folgend werden für Bareinnahmen mit der Pflicht zur taggenauen Erfassung von Kasseneinnahmen und -ausgaben¹⁴² sowie neuerdings dem Schutz eines elektronischen Aufzeichnungssystems und der digitalen Aufzeichnungen durch eine zertifizierte technische Sicherheitseinrichtung¹⁴³ deutlich strengere Ordnungsvorgaben angelegt. Entsprechend infiziert deren Nichteinhaltung nach ständiger Rechtsprechung die gesamte Buchführung.¹⁴⁴

Der **Anwendungsbereich** der Ordnungsvorschriften wird in den GoBD angegeben mit: „Somit sind alle Unternehmensbereiche betroffen, in denen betriebliche Abläufe durch

¹³⁶ Vgl. z.B.: BFH v. 2.10.1968, BStBl. II 1969, 157; Zeitgerechte Erfassung von unbaren Geschäftsvorfällen innerhalb von zehn Tagen.

¹³⁷ Insbesondere: „Grundsätze zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)“, BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450.

¹³⁸ FG Hamburg v. 26.8.2016 – 6 V 81/16, 1. Leitsatz.

¹³⁹ Vgl. z.B.: BFH v. 16.12.2014, BStBl. II 2015, 519, 2. Leitsatz: „Verwendet ein Einzelhändler, der in seinem Betrieb im allgemeinen Waren von geringem Wert an ihm der Person nach nicht bekannte Kunden über den Ladentisch gegen Barzahlung verkauft, eine PC-Kasse, die detaillierte Informationen zu den einzelnen Verkäufen aufzeichnet und eine dauerhafte Speicherung ermöglicht, so sind die damit bewirkten Einzelaufzeichnungen auch zumutbar“ sowie BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27: „Der erkennende Senat vertritt hierzu die Auffassung, dass das Fehlen einer lückenlosen Dokumentation zur Kassenprogrammierung in seinen Auswirkungen auf die Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis dem Fehlen von Tagesendsummenbons bei einer Registrierkasse bzw. dem Fehlen täglicher Protokolle über das Auszählen einer offenen Ladenkasse gleichsteht. In allen drei Fällen lässt der formelle Mangel zwar keinen sicheren Schluss auf die Verkürzung von Einnahmen zu. Gleichwohl gibt es systembedingt keine Gewähr mehr für die Vollständigkeit der Erfassung der Bareinnahmen, ohne dass eine nachträgliche Ergänzung der Dokumentation bzw. eine anderweitige Heilung des Mangels möglich wäre.“

¹⁴⁰ Vgl. a.: „Neufassung des § 146 Abs. 1 AO durch das Gesetz zum Schutz vor Manipulationen an digitalen Grundaufzeichnungen vom 22. Dezember 2016; Anwendungserlass zu § 146 AO“, BMF-Schreiben v. 19.6.2018, Pkt. 2.1.7.

¹⁴¹ Vgl. z.B.: BFH v. 7.2.2008 – X B 189/07: „Allerdings sind auch Steuerpflichtige, die den Gewinn nach § 4 Abs. 3 EStG ermitteln, verpflichtet, die ihrer Gewinnermittlung zugrunde liegenden Belege aufzubewahren. [...] Zutreffend hat das FG darauf abgestellt, dass gerade in Fällen, in denen Steuerpflichtigen eine Einzelaufzeichnungspflicht nicht zugemutet werden kann, die Einnahmeermittlung – z.B. bei Einsatz von Registrierkassen durch Erstellung und Aufbewahrung der Kassensummenbons – nachvollziehbar dokumentiert und überprüfbar sein muss.“

¹⁴² § 146 Abs. 1 AO.

¹⁴³ § 146a Abs. 1 AO; eingefügt durch das Gesetz zum Schutz vor Manipulationen an digitalen Grundaufzeichnungen vom 22.12.2016 und anzuwenden ab 1.1.2020.

¹⁴⁴ Vgl. z.B.: BFH v. 14.12.2011 – XI R 5/10 oder FG Sachsen v. 4.4.2008 – 5 V 1035/07.

DV-gestützte Verfahren abgebildet werden und ein Datenverarbeitungssystem [...] für die Erfüllung der [...] außersteuerlichen oder steuerlichen Buchführungs-, Aufzeichnungs- oder Aufbewahrungspflichten verwendet wird. [...] Die Erfassung der Geschäftsvorfälle in elektronischen Grund(buch)aufzeichnungen [...] und die Verbuchung im Journal [...] kann organisatorisch und zeitlich auseinander fallen [...]. In diesen Fällen gelten die Ordnungsvorschriften bereits mit der ersten Erfassung [...] und müssen über alle nachfolgenden Prozesse erhalten bleiben¹⁴⁵. Diese Reichweite der Ordnungsmaßgaben findet sich regelmäßig in der Rechtsprechung¹⁴⁶ wieder und ist erforderlich, weil sich die Buchführung beim Vorliegen detaillierter Grundaufzeichnungen i.d.R. auf diese bezieht und alleine nicht den Formalanforderungen entspricht. Besonders trifft dieser erweiterte Anwendungsbereich der Ordnungsvorschriften auf Kassenslösungen zur Erstaufzeichnung zu, weil üblicherweise keine weiteren Nachweise der Geschäftsvorfälle vorliegen und das hohe Manipulationsrisiko bekannt ist¹⁴⁷. Folglich wird in der Literatur die Beraterverantwortung für die Kassensführung hervorgehoben.¹⁴⁸

Die eine zentrale Frage der Formalprüfung ist: „Lässt sich die **Vollständigkeit der Aufzeichnungen nachvollziehen?**“ Nach den Ordnungsvorschriften der Reichsabgabenordnung sollte dies durch fest gebundene Bücher mit durchnummerierten Seiten (weitgehend) sichergestellt werden.¹⁴⁹

Als Folge der Verbreitung elektronischer Betriebsverwaltungen und der damit verbundenen Möglichkeiten und Gefahren (s. [Pkt. 2.1.](#)) ist durch das Gesetz zum Schutz vor Manipulationen an digitalen Grundaufzeichnungen vom 22.12.2016 die Pflicht zur Einzelaufzeichnung explizit in § 146 Abs. 1 AO aufgenommen worden. Ein mögliches Kriterium zur Überprüfung der vollständigen Einzelaufzeichnung in digitalen Betriebsdaten ist die programmsei-

¹⁴⁵ BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450, Rz. 9 u. 84.

¹⁴⁶ Vgl. z.B.: BFH v. 14.12.1966, BStBl. III 1967, 247: „Eine körperliche Bestandsaufnahme gehört in der Regel ebenfalls zu den unabdingbaren Voraussetzungen der Ordnungsmäßigkeit einer Buchführung. Fehlt eine zureichende körperliche Bestandsaufnahme, so ist die Buchführung mangelhaft, auch wenn das FA die Bestände – vielleicht zutreffend – schätzt oder die Schätzung des Steuerpflichtigen unverändert übernimmt“ sowie BFH. V. 26.2.2004, BStBl. II 2004, 599: „Die Schichtzettel sind Einnahmehauptaufzeichnungen; sie enthalten Angaben, aus denen sich die Höhe der Umsätze und damit auch der Betriebseinnahmen unmittelbar ergibt. Der Kläger war insoweit aufbewahrungspflichtig; die Vernichtung der Schichtzettel bedeutete einen Verstoß gegen diese Pflicht.“

¹⁴⁷ Vgl. z.B.: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27: „Elektronische Kassensysteme sind durch Umprogrammierung in nahezu beliebiger Weise manipulierbar; von derartigen Manipulationsmöglichkeiten machen Teile der betrieblichen Praxis nach dem Erkenntnisstand des Senats durchaus Gebrauch (zu einem solchen Fall z.B. Beschluss des FG Rheinland-Pfalz vom 7. Januar 2015 5 V 2068/14; vgl. zum Ganzen auch Tz. 54 der Bemerkungen des Bundesrechnungshofs 2003 zur Haushalts- und Wirtschaftsführung, BTDrucks 15/2020, 197 f.).“

¹⁴⁸ Pump/Heidl, „Hilfestellung bei der Kassensführung durch den steuerlichen Berater“, StBp 2014 S 162, 163: „Unverantwortliche Sorglosigkeit kann die Berater dazu verleiten, sich nicht um die Kassensführung ihrer Mandanten zu kümmern. Das wird damit begründet, dass dafür kein Prüfungsauftrag bestehe. Wer am Fortbestand seiner Mandanten interessiert ist, muss sie vor fehlerhafter Kassensführung schützen.“

¹⁴⁹ § 162 Abs. 4 RAO.

tig mitgeschriebene Sequenznummer: „Die vollständige und lückenlose Erfassung und Wiedergabe aller Geschäftsvorfälle ist bei DV-Systemen durch ein Zusammenspiel von technischen [...] und organisatorischen Kontrollen sicherzustellen (z.B. [...] automatisierte Vergabe von Datensatznummern“¹⁵⁰.

Auch die regelmäßig als Grund für eine Ausnahme von der Einzelaufzeichnungspflicht angeführten Abwägungen des BFH zur Zumutbarkeit bei Barverkäufen geringen Werts an eine Vielzahl von (überwiegend unbekannt) Kunden¹⁵¹ werden inzwischen im Kontext der heutzutage üblichen digitalen Aufzeichnungs- und Aufbewahrungsweise bewertet: „Aus Gründen der Zumutbarkeit und Praktikabilität hat der BFH in der Vergangenheit eine Pflicht zur Einzelaufzeichnung für den Einzelhandel und vergleichbare Berufsgruppen verneint ([...] z.B. bei einer Vielzahl von einzelnen Geschäften mit geringem Wert, Verkauf von Waren an der Person nach unbekannte Kunden über den Ladentisch gegen Barzahlung). Werden elektronische Grund(buch)aufzeichnungen [...] tatsächlich [...] geführt, dann sind diese Daten auch aufzubewahren und in maschinell auswertbarer Form vorzulegen. Insoweit stellt sich die Frage der Zumutbarkeit und Praktikabilität nicht.“¹⁵². Der BFH bestätigt dieses Abstellen auf die faktischen Verhältnisse des Einzelfalls.¹⁵³

Die andere zentrale Frage der Formalprüfung ist: „Lässt sich die Unverändertheit der Aufzeichnungen nachvollziehen oder zumindest erkennen, was verändert wurde und wie der ursprüngliche Zustand war (**Gewissheit der Originalität**)?“ Hierzu wurde in der Reichsabgabenordnung die Verwendung des unauslöschlichen Schreibmaterials „Tinte“ verlangt.¹⁵⁴

Derweil der Grundzustand elektronisch erfasster Informationen „veränderbar“ ist¹⁵⁵, erfordert die Gewissheit der Originalität aktive Maßnahmen: „Das zum Einsatz kommende DV-Verfahren muss die Gewähr dafür bieten, dass alle Information (Programme und Datenbe-

¹⁵⁰ BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450, Rz. 40.

¹⁵¹ Urte. v. 12.5.1966, BStBl. III 1966, 371.

¹⁵² BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450, Rz. 39.

¹⁵³ BFH v. 16.12.2014, BStBl. II 2015, 519, Rz. 20: „Da die GoB indes nur eine Einzelaufzeichnung der Kassenvorgänge im Rahmen des nach Art und Umfang des Geschäftes Zumutbaren verlangen, hat der BFH die Einzelaufzeichnungspflicht für Einzelhandelsgeschäfte – in Betrieben, in denen Waren von geringerem Wert an eine unbestimmte Vielzahl nicht bekannter und auch nicht feststellbarer Personen verkauft werden – dahingehend eingeschränkt, dass die baren Betriebseinnahmen in der Regel nicht einzeln aufgezeichnet zu werden brauchen (BFH-Urteil in BFHE 86, 118, BStBl. III 1966, 371). Ausschlaggebend für den BFH war insoweit, dass es technisch, betriebswirtschaftlich und praktisch unmöglich war, an die Aufzeichnung der einzelnen zahlreichen baren Kassenvorgänge in Einzelhandelsgeschäften gleiche Anforderungen wie bei anderen Handelsgeschäften zu stellen, nämlich zur Identifizierung und zur Bestimmung des Inhalts des Geschäfts Namen und Anschrift des Kunden und den Gegenstand des Kaufvertrages festzuhalten.“

¹⁵⁴ § 162 Abs. 6 RAO.

¹⁵⁵ Vgl. z.B.: Tipke/Kruse, AO/FGO, 141. Lieferung 7/2015, § 146 AO, Rz. 59: „Moderne EDV-Buchführungssysteme [...] müssen programmmäßige Sicherungen und Sperren enthalten, die verhindern, dass einmal eingegeben Daten dem Zugriff zur Wiederaufbereitung (Änderung) preisgegeben sind. [...] Dies muss bereits vom Zeitpunkt der erstmaligen Speicherung an, nicht erst nach durchgeführter Verarbeitung gewährleistet sein“.

stände), die einmal in den Verarbeitungsprozess eingeführt werden (Beleg, Grundaufzeichnung, Buchung), nicht mehr unterdrückt oder ohne Kenntlichmachung überschrieben, gelöscht, geändert oder verfälscht werden können. [...] Die Unveränderbarkeit der Daten, Datensätze, elektronischen Dokumente und elektronischen Unterlagen [...] kann sowohl hardwaremäßig (z.B. unveränderbare und fälschungssichere Datenträger) als auch softwaremäßig (z.B. Sicherungen, Sperrungen, Festschreibung, Löschermerker, automatische Protokollierung, Historisierungen, Versionierungen) als auch organisatorisch (z.B. mittels Zugriffsberechtigungskonzepten) gewährleistet werden. Die Ablage von Daten und elektronischen Dokumenten in einem Dateisystem erfüllt die Anforderungen der Unveränderbarkeit regelmäßig nicht, soweit nicht zusätzliche Maßnahmen ergriffen werden, die eine Unveränderbarkeit gewährleisten.¹⁵⁶ Die jüngere Gesetzesaktualisierung zum Schutz eines elektronischen Aufzeichnungssystems und der digitalen Aufzeichnungen durch eine zertifizierte technische Sicherheitseinrichtung¹⁵⁷ dienen im Wesentlichen der Sicherstellung der Originalität von Besteuerungsdaten.

In der Literatur wird inzwischen zunehmend die „Ordnungsmäßigkeitsberatung“ als Möglichkeit gesehen, sich innerhalb der Beraterschaft durch ausreichende Kenntnisse der heutzutage üblichen komplexen Aufzeichnungs- und Aufbewahrungsprozesse hervorzuheben.¹⁵⁸

Z - Bericht		
04.01.2008 - 00:58		
Nr. 1		
Bereich: Alle Kassen		
Von: 03.01.2008 - 14:29		
Bis: 04.01.2008 - 00:40		
Brutto total	335	880.60
Netto total	335	740.00
Retoure		
Master Storno		
Storno		
Umsatz Brutto MwSt 19.00%		880.60
Umsatz Netto MwSt 19.00%		740.00
MwSt 19.00%		140.60
Prozentanteil		100.00
Durchschn. MWST Satz		19.00
Zahlungsarten:		
in Bar	335	880.60
Vorgangsdurchschnitt	96	9.17
Kundendurchschnitt	96	9.17
Umsatz		880.60

Der Umgang mit den Mindestanforderungen an die vereinfachte Dokumentation von elektronischen Registrierkassen mittels Tagesendsummenbons (Z-Bons) lt. Erlass¹⁵⁹ hat jahrelang einen Teil der Prüfungspraxis mitbestimmt. So leidet die Gewissheit der Originalität im abgebildeten Z-Bon-Beispiel¹⁶⁰ an der fehlenden Ausgabe von Korrekturen. Da es beim Verkauf einer Vielzahl von Leistungen (bzw. Waren) äußerst unwahrscheinlich ist, dass – trotz rechtzeitiger Aufzeichnung – keine Erfassungsfehler geschehen, muss der nachteilige Schluss¹⁶¹ gezogen wer-

¹⁵⁶ BMF-Schreiben v. 14.11.2014 BStBl. I 2014, 1450, Rz. 108 u. 110.

¹⁵⁷ § 146a Ab. 1 AO; eingefügt durch das Gesetz zum Schutz vor Manipulationen an digitalen Grundaufzeichnungen vom 22.12.2016 und anzuwenden ab 1.1.2020.

¹⁵⁸ Vgl. z.B.: *Liekenbrock*, „Digital Tax Compliance – Praxishinweise zur technischen Umsetzung von Tax Compliance Management Systemen“, Ubg 2018, 43.

¹⁵⁹ „Elektronische Registrierkassen, Kassenstreifen“, BFM-Schreiben v. 9.1.1996, BStBl. I 1996, 34.

¹⁶⁰ Bildquelle: *Wähnert*, „Die Beweislast im Besteuerungsverfahren – Brennpunkt ‚Einnahmenvollständigkeit‘“, StC 11/2011, 16, 17.

¹⁶¹ BFH v. 15.2.1989, BStBl. II 1989, 462: „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO) [...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

den, dass Einnahmenverkürzungen verschleiert werden. Dieser Eindruck wird noch dadurch bekräftigt, dass die drei Stornoebenen nicht die Tagessumme „0“ anzeigen, sondern offensichtlich in den Berichtseinstellungen deaktiviert wurden.

Eine „funktionale“¹⁶² **Übertragung der Hauptordnungskriterien** „nachvollziehbare Vollständigkeit“ und „Gewissheit der Originalität“ auf ein zeitgemäßeres Aufzeichnungssystem mit dauerhafter Speichermöglichkeit bedeutet im Fall des nebenstehenden Auszugs aus den Kassendaten, dass die **Lücke in den fortlaufenden Vorgangsnummern „RENr“** folgende Ursachen haben kann:

1. Der Sachverhalt zur RENr „7522“ ist aufgrund einer bestimmten Einstellung des Kassensystems in dem Datenexport nicht mit ausgegeben worden.
2. Durch einen Eingriff in die Aufzeichnung der Verkaufsvorgänge ist der eine mit der RENr „7522“ gelöscht worden.
3. Es sind mehrere Vorgänge nachträglich gelöscht worden. Bei der Neunummerierung der Datensätze ist jedoch die RENr „7522“ vergessen worden.

DATUM	ZEIT	RENr	TOTAL	BEDNR	KASSE
22.12.2007	23:31	7503	2,50	5	1
22.12.2007	23:40	7504	15,00	5	1
22.12.2007	23:42	7505	7,50	5	1
22.12.2007	23:42	7506	5,00	5	1
22.12.2007	23:44	7507	2,50	5	1
22.12.2007	23:46	7508	3,00	5	1
22.12.2007	23:49	7509	7,00	5	1
22.12.2007	23:50	7510	2,20	5	1
22.12.2007	23:51	7511	2,50	5	1
22.12.2007	23:52	7512	4,50	5	1
22.12.2007	23:55	7513	9,00	5	1
23.12.2007	00:00	7514	10,00	5	1
23.12.2007	00:02	7515	3,50	5	1
23.12.2007	00:05	7516	3,00	5	1
23.12.2007	00:07	7517	9,70	5	1
23.12.2007	00:08	7518	6,90	5	1
23.12.2007	00:08	7519	2,50	5	1
23.12.2007	00:08	7520	2,50	5	1
23.12.2007	00:09	7521	2,50	5	1
23.12.2007	00:14	7523	5,00	5	1
23.12.2007	00:16	7524	2,20	5	1
23.12.2007	00:16	7525	7,50	5	1

Weil der tatsächliche Grund für die fehlerhafte Datenausgabe ohne die vollständigen Systeminformationen inkl. aller Einstellungsprotokolle nicht ermittelbar ist, sieht der BFH in dem „Fehlen einer lückenlosen Dokumentation zur Kassenprogrammierung in seinen Auswirkungen auf die Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis [einen Ordnungsmangel, der] dem Fehlen von Tagesendsummenbons bei einer Registrierkasse bzw. dem Fehlen täglicher Protokolle über das Auszählen einer offenen Ladenkasse gleichsteht.“¹⁶³

Obwohl die beiden Voraussetzungen der „Beweiskraft der Buchführung (und Aufzeichnungen)“ nach § 158 AO „Ordnungsmäßigkeit“ und „kein Beanstandungsanlass“ lt. Gesetzestext voneinander unabhängig und gleichwertig sind¹⁶⁴, erfährt die Formalprüfung in der aktuellen

¹⁶² FG Hamburg v. 26.8.2016 – 6 V 81/16, 1. Leitsatz.

¹⁶³ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

¹⁶⁴ § 158 AO „Beweiskraft der Buchführung: Die Buchführung und die Aufzeichnungen des Steuerpflichtigen, die den Vorschriften der §§ 140 bis 148 entsprechen, sind der Besteuerung zugrunde zu legen, soweit nach den Umständen des Einzelfalls kein Anlass ist, ihre sachliche Richtigkeit zu beanstanden.“

Rechtsprechung eine zunehmend zeitgemäßere Auslegung¹⁶⁵, während das für die Finanzverwaltung abgesenkte Gewissheitsmaß nach wie vor auf der Grundlage eines höchsttrichterlich geforderten Übermaßes¹⁶⁶ (s. [Pkt. 4.6.](#)) umgesetzt und somit faktisch konterkariert wird. Der Zusammenhang zwischen formellen und inhaltlichen Mängeln von Besteuerungsaufzeichnungen wird ungeachtet dessen gesehen.¹⁶⁷ Deshalb führt „[d]ie Vorlage einer bereits formell ordnungswidrigen Buchführung [...] von vornherein zu einer erheblichen Absenkung des Beweismaßes für eine Schätzung“¹⁶⁸.

Deshalb ist es bei Beanstandungen der sachlichen Richtigkeit von Besteuerungsgrundlagen – insbesondere im Zusammenhang mit Gesamtbildprüfungen¹⁶⁹ – wichtig, vorgefundene **Ordnungsmängel ausführlich zu dokumentieren** wie in diesem Beispiel: „Das beklagte Finanzamt (FA) führte [...] beim Kläger eine Außenprüfung [...] durch, in deren Verlauf es zu der Einschätzung kam, die Kassenbuchführung des Klägers sei nicht ordnungsgemäß. Die Prüferin hat im Verlauf der Prüfung diesbezüglich folgende Feststellungen getroffen (...):

- die Kasseneinnahmen wurden zum Teil nicht täglich festgehalten, sondern in einem Z-Bon, gleich für mehrere Tage, zusammengefasst,

¹⁶⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27: „Der erkennende Senat vertritt hierzu die Auffassung, dass das Fehlen einer lückenlosen Dokumentation zur Kassenprogrammierung in seinen Auswirkungen auf die Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis dem Fehlen von Tagesendsummenbons bei einer Registrierkasse bzw. dem Fehlen täglicher Protokolle über das Auszählen einer offenen Ladenkasse gleichsteht. In allen drei Fällen lässt der formelle Mangel zwar keinen sicheren Schluss auf die Verkürzung von Einnahmen zu. Gleichwohl gibt es systembedingt keine Gewähr mehr für die Vollständigkeit der Erfassung der Bareinnahmen, ohne dass eine nachträgliche Ergänzung der Dokumentation bzw. eine anderweitige Heilung des Mangels möglich wäre.“

¹⁶⁶ BFH v. 24.6.1997, BStBl. II 1998, 51: „Nach § 158 AO 1977 sind der Besteuerung die Buchführung und die Aufzeichnungen des Steuerpflichtigen, die den Vorschriften der §§ 140 bis 148 AO 1977 entsprechen, zugrunde zu legen, soweit nach den Umständen des Einzelfalls kein Anlaß besteht, ihre sachliche Richtigkeit zu beanstanden. Nur wenn die Würdigung des Sachverhalts ergibt, daß eine formell ordnungsmäßige Buchführung mit an Sicherheit grenzender Wahrscheinlichkeit sachlich unrichtig ist, kann das Ergebnis der Buchführung ganz oder teilweise verworfen werden.“

¹⁶⁷ Vgl. z.B.: Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist“ oder *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Strafverfahren“, NZWiSt 1/2014, 24, 25: „In der Praxis wird sich die Frage, ob ein Zeitreihenvergleich allein die Beweisvermutung des § 158 AO widerlegt, [...] dadurch relativieren, dass die [...] aufgedeckten Unregelmäßigkeiten regelmäßig mit Unregelmäßigkeiten in der Kassenbuchführung einhergehen.“

¹⁶⁸ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9.

¹⁶⁹ Gesamtbildprüfungen (Verprobungen) können grds. die vom BFH geforderte Aussagegewissheit „mit an Sicherheit grenzender Wahrscheinlichkeit“ nicht erreichen, weil sie komplexe Sachverhalte auf Grundlage von unvollständigen, evtl. nicht objektiven Informationen aus der Vergangenheit wiederherstellen (s.a. [Pkt. 4.6.](#)). Mehrperspektivisch systematische Prüfungsnetze wie die SRP erreichen zwar über die Absicherung von Einzelergebnissen und über Indiziengesamtheiten ein qualifizierteres Gesamtbild, sind aber dennoch wie alle Prüfungsmaßnahmen von der Datenqualität abhängig.

- die Z-Zähler Nr. fängt mehrfach wieder mit Nr. 1 an, und zwar in 2010 am: 08.03., 12.03., 10.07., 04.08. und 21.12., in 2011 am: 26.02. und in 2012 am: 15.08., wobei dies auf die Anschaffung der neuen Kasse zurückzuführen ist,
- die Erlöse aus dem Z-Bon Nr. 160 vom 01.09.2011 wurden nicht als Einnahme erfasst (...),
- der Z-Bon Nr. 172 im September 2011 fehlt (...),
- die Erlöse für Sonntag, den 22.01.2012, fehlen (...),
- in allen Streitjahren wurden zum Teil hohe Stornobuchungen vorgenommen (...),
- für den Zeitraum 13.01.2010 bis 25.01.2010 liegt lediglich der Z-Bon mit der Nr. 97 vor (... €), (...),
- für den Zeitraum vom 08.06.2010 bis zum 06.07.2010 liegt nur ein Z-Bon vom 07.07.2010 mit der Nr. 84 vor, wobei der darin ausgewiesene Betrag (... €) in den monatlichen Exceltabellen aufgeteilt wurde in Höhe von ... € auf den 30.06. und in Höhe von ... € auf den 06.07. (...),
- für den Zeitraum vom 07.01.2011 bis zum 14.02.2011 liegt lediglich ein Z-Bon mit der Nr. 16 (Einnahme: ... €) vor,
- für den Zeitraum 15.02.2011 bis 25.02.2011 liegen keine Z-Bons vor (...), ohne dass am 26.02.2011 ein entsprechend höherer Betrag erfasst worden wäre (... €), (...),
- für den 12.01.2010 (...), 12.08.2012 (...) und den 21.08.2012 (...) liegen nur handschriftlich gefertigte Aufzeichnungen über die Tageseinnahmen vor,
- für den Zeitraum 16.09.2011 bis 29.09.2011 wurden keine Tageserlöse erklärt, dann am 30.09.2011 ein Betrag in Höhe von ... € (Z-Bon 174; ...),
- für den Zeitraum 01.10.2011 - 05.10.2011 wurden ebenfalls keine Tageserlöse erfasst, ohne dass aber am 06.10.2011 ein entsprechend höherer Betrag erfasst worden wäre (... € vgl. Z-Bon 175; ...) und
- in 2012 wurden Kasseneinnahmen nur für 302 Tage erklärt, so dass bei einem Abzug von 52 Ruhetagen für 12 Tage keine Einnahmen erklärt wurden (...).¹⁷⁰

Außerdem sollten konkrete Bedenken hinsichtlich der inhaltlichen Richtigkeit der Besteuerungsgrundlagen aus diesen Formalmängeln abgeleitet werden, weil zuletzt wiederholt betont worden ist, dass die Ordnungsmäßigkeitsprüfung nicht dem Selbstzweck dient¹⁷¹, son-

¹⁷⁰ FG Hamburg v. 5.3.2018 – 3 K 205/15, Rz. 14.

¹⁷¹ Vgl. z.B.: Nöcker (RiBFH), „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3550: „Wie es § 158 i. V. mit § 162 Abs. 2 Satz 2 zweite Alternative AO bereits von seinem Wortlaut her vorsieht, ist eine Schätzung dem Grunde nach bei einem Steuerpflichtigen, der eine Buchführung (freiwillig oder gesetzlich verpflichtend) führt, vorzunehmen, wenn die Vorschriften der §§ 140 bis 148 AO verletzt sind. Die ständige höchstrichterliche Rechtsprechung schränkt diese allein auf formellen Buchführungsmängeln beruhende Schätzungsbezugnis aber insoweit ein, als sie allein Anlass geben, die sachliche Richtigkeit der Buchführungsergebnisse anzuzweifeln.“

dem die Prüfbarkeit durch einen „sachverständigen Dritten innerhalb angemessener Zeit“¹⁷² sicherstellen soll.

4.2. Regelbasierte Ansätze (Zeitreihenanalysen)

Regelbasierte Prüfungsansätze sollen entweder die **Übereinstimmung** zweier Datenreihen auf chronologischer Ebene nachverfolgen (*s. das Beispiel der Schnittstellenprüfung unter [Pkt. 5.3.](#)*) **oder** die **Plausibilität** von (Betriebs-)Daten anhand der betriebswirtschaftlich erwarteten Abhängigkeit beurteilen. Gesamtbildverprobungen, die sich auf die (kaufmännische) Schlüssigkeit gründen, gehen bis auf die Reichsfinanzhof-Rechtsprechung¹⁷³ zurück und sind nach dem Verifikationsprinzip¹⁷⁴ notwendig, weil „es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann“¹⁷⁵.

Anders als die im Revisionsbereich verbreiteten Zeitreihenanalysen zur Plausibilitätsprüfung, die sich auf die Entwicklung der Verhältnisgröße „Aufschlagsatz“ beziehen¹⁷⁶, setzt die SRP zentral auf einen **doppelskalierten Zeitreihenvergleich** der zwei auf einen Zusammenhang zu prüfenden Datenreihen, denn „[b]ei dieser Methode tritt das zahlenmäßige Verhältnis zu Gunsten der Abhängigkeit voneinander in den Hintergrund. Wesentlich ist die Reaktion der Faktoren aufeinander.“¹⁷⁷ Während die Entwicklung des Aufschlagsatzes im Unternehmensalltag vielen Einflüssen ausgesetzt ist und deshalb in der Literatur entsprechende Unwegbarkeiten aufgelistet werden¹⁷⁸, gibt die zentrale SRP-Zeitreihe die Handlungslogik im Betrieb erheblich eindeutiger wieder. Somit erreichen Schlüssigkeitsurteile gem. § 158 AO eine deutlich höhere Gewissheit.

Voraussetzung für die Auswahl **geeigneter Methoden** zur jeweiligen Prüfungsfrage ist eine ausreichende **Methodenkenntnis**, welche – neben der systematisch mehrperspektivischen Vernetzung – einen Fokus der SRP bildet (*s.a. den Abschnitt „Methodenwissen“ unter*

¹⁷² § 145 Abs. 1 AO.

¹⁷³ Vgl. z.B.: RFH v. 13.1.1937, RStBl 1937, 317.

¹⁷⁴ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungen wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

¹⁷⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5.

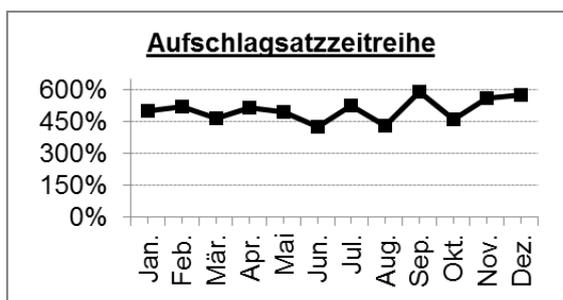
¹⁷⁶ Vgl. z.B.: FG Münster v. 19.8.2004 – 8 V 3055/04, FG Düsseldorf v. 15.2.2007 – 16 V 4691/06 oder FG Köln v. 27.1.2009 – 6 K 3954/07. Auch das BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743) bezieht sich mit der verwendeten GleitschlittENZEITREIHE auf eine Zeitreihenvariante des Aufschlagsatzes.

¹⁷⁷ FG Hamburg v. 31.10.2016 – 2 V 202/16, Rz. 48.

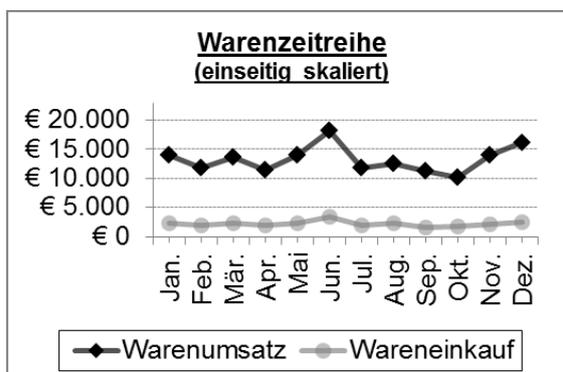
¹⁷⁸ Vgl. z.B.: Wiggen, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168.

Pkt. 2.2.). Lt. BFH „sind zahlreiche Varianten eines Zeitreihenvergleichs denkbar“¹⁷⁹. Im Gegensatz zum FG Hamburg¹⁸⁰ haben sich die RiBFH aber nicht mit den Unterschieden auseinandergesetzt, sondern Aufschlagsatzzeitreihen stellvertretend für alle Formen beurteilt¹⁸¹. Weil sich dieses Vorgehen bei Kenntnis der eklatanten Modellunterschiede verboten hätte, sollen hier drei Varianten der Zeitreihenanalyse zu denselben Daten gegenübergestellt werden:

1. Das erste Diagramm zeigt die monatliche Entwicklung des Aufschlagsatzes aus Wareneinkauf und -umsatz für eine kleine Kneipe, die nach § 4 Abs. 3 EStG aufzeichnet. Dabei ist eine Schwankungsbreite des Erfolgsfaktors von unter 450 % bis fast 600 % zu beobachten. Für die Prüfungsfrage der Gesamtschlüssigkeit ist beurteilungserheblich, ob diese „Regelstörungen“¹⁸² auf Zuordnungsunschärfen des Wareneinkaufs bzw. dessen Bezahlung, auf betriebliches Verhalten wie bspw. Lagerankäufe größerer Getränkemengen mit längerer Haltbarkeit oder auf sachlich falsche Besteuerungsdaten zurückgehen.



2. In der nächsten Auswertung werden die – der ersten Analyse zu Grunde liegenden – beiden Datenreihen zum monatlichen Wareneinkauf und -umsatz als Einzelgrafiken mit einer Y-Achse¹⁸³ aufgetragen. Hierbei werden die Absatzveränderungen mit starken Anstiegen im Sommer und Winter ersichtlich. Die Veränderungen im Wareneinkauf erscheinen deutlich moderater. Es ist aber zumindest erkennbar, dass der verstärkte Verkauf in den beiden Hochsaison-Zeiträumen auch eine Erhöhung des Einkaufs nach sich zieht.



3. Um die unternehmerische Handlungslogik eines gesteuerten Wareneinkaufs in Abhängigkeit vom aufkotroyierten Absatz in der Datenvisualisierung nachvollziehbar wiederzugeben, werden die beiden Datenreihen in der dritten Grafik mit jeweils einer eigenen Y-

¹⁷⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 37.

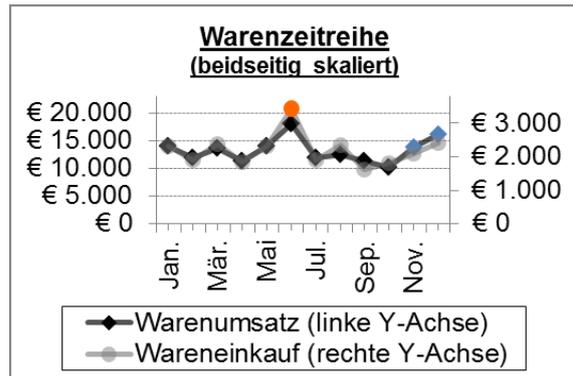
¹⁸⁰ FG Hamburg v. 31.10.2016 – 2 V 202/16, Rz. 48.

¹⁸¹ Auch die im BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743) verhandelte Gleitschlitzenzeitreihe bezieht sich auf eine chronologische Nachverfolgung des Aufschlagsatzes.

¹⁸² Die Prüfungsregel bei dieser Datenverbildlichung ergibt sich mit der (weitgehend) erwarteten Konstanz der Verhältnisgröße „Aufschlagsatz“.

¹⁸³ Der einseitig skalierte Zeitreihenvergleich ist in diversen Datenanalyseprogrammen (zurzeit noch) die Standardauswertungsform.

Achse¹⁸⁴ abgebildet. Infolge der Größenanpassung des Wareneinkaufs an den Warenumsatz lässt sich bei dieser Zeitreihenanalyse unzweifelhaft ablesen, dass grundsätzlich ein enger Zusammenhang der beiden Betriebsfaktoren besteht. Im Zeitraum „August – September“ liegt offenbar tatsächlich eine Zuordnungsunschärfe



vor, indem der erhöhte Wareneinkauf des August einen geringeren im nächsten Monat zur Folge hat. Neben dem insgesamt plausiblen Zusammenhang zwischen Marktfaktor „Absatz“ und unternehmerischem Verhalten „Wareneinkauf“ lassen sich der Auswertung zudem zwei klare Vertiefungsbereiche entnehmen: Erstens ist der **Wareneinkauf im Juni** (relativ) zu hoch, was als punktuell Pruffeld untersucht werden kann. Zweitens übertrifft der **Umsatz ab November** den für Einkauf dauerhaft. Hinter dieser Entwicklung ist eine Verkaufspreiserhöhung zu vermuten, die – nach Verifizierung – in der weiteren Gesamtbildprüfung zu beachten ist.

Da drei Zeitreihenvarianten zur Plausibilitätsprüfung derselben Warenflussdaten zu (stark) abweichenden Ergebnissen führen, muss die jeweilige Eignung mit Hilfe einer objektiven Gegenkontrolle hinterfragt werden. Die statistische Kenngröße für die gleichgerichtete Abhängigkeit zweier Datenreihen ist der **Korrelationskoeffizient**¹⁸⁵. Zum Beispielsfall resultiert mit dem Wert „+0,90“ eine sehr gute positive Gleichrichtung. In der Konsequenz sind die ersten beiden Zeitreihentypen für die Plausibilitätsprüfung „Sind zwei Betriebsgrößen – wie erwartet – voneinander abhängig?“ nur begrenzt geeignet, während insbesondere der doppelskaliert größenangepasste Zeitreihenvergleich als eigenständiger Ansatz gewertet werden muss.

Deshalb wird i.R. der SRP der Einsatz von Zeitreihenanalysen streng nach dem **Verwendungszweck** differenziert:

- ⊙ Soll die Entwicklung einer Betriebsgröße chronologisch nachverfolgt werden, geschieht dies durch eine Einzelgrößenzeitreihe wie beim Aufschlagsatz.
- ⊙ Sollen zwei Datenreihen wie bspw. bei einer Schnittstellenprüfung (s. [Pkt. 5.3.](#)) auf Übereinstimmung geprüft werden, hat dies in einer einzelskalierten Gegenüberstellung

¹⁸⁴ Tabellenkalkulationsprogramme wie z.B. MS Excel® stellen die größenanpassende Doppelskalierung als Standardoption für Zeitreihenanalysen zur Verfügung.

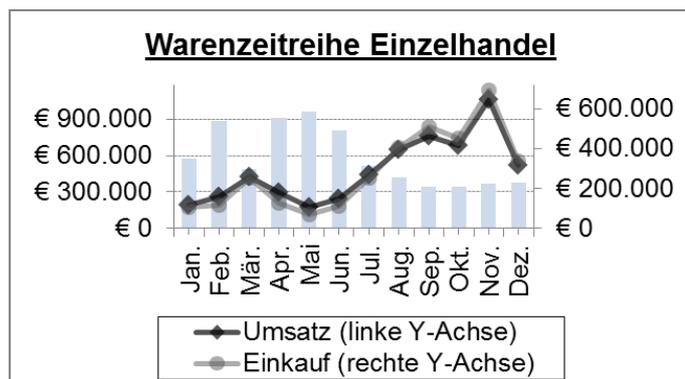
¹⁸⁵ Als Maß für den Zusammenhang zweier Datenreihen nimmt der Korrelationskoeffizient Größen von „-1“ für die 100 %-ige Gegenläufigkeit der Grafen über „0“ bei nicht vorhandener Abhängigkeit bis zu „+1“ für eine vollständig positive Gleichrichtung ein.

zu erfolgen, damit der Vergleich nicht durch eine automatische Größenanpassung verzerrt wird.

- © Soll z.B. mit dem Warenfluss die (chronologische) Abhängigkeit zweier Unternehmensfaktoren als logischer Handlungszusammenhang untersucht werden, erfolgt dies über einen doppelskaliert größenangepassten Zeitreihenvergleich.

Wegen der sehr guten Eignung des doppelskaliert größenangepassten Zeitreihenvergleichs für die Frage der Abhängigkeit zweier Datenreihen ist diese Zeitreihenform seit Beginn der SRP das zentrale Modell für die Plausibilitätsprüfung gem. § 158 AO.¹⁸⁶ Auf diese Weise wird bspw. bzgl. der Warenflussdaten untersucht, ob der Warennachkauf (möglichst) eng an dem vom Markt vorgegebenen Warenumsatz orientiert wird. Der Schlüssigkeitsansatz ist deutlich zwingender als die chronologische Nachverfolgung des Aufschlagsatzes, weil dieser in seiner Konstanz erheblich von üblichen Betriebseinflüssen wie z.B. wechselnden Mengenverhältnissen mit unterschiedlichen Aufschlagsätzen oder konjunkturellen Preisveränderungen bestimmter Produkte bzw. Leistungen gestört werden kann¹⁸⁷.

Die **Robustheit des doppelskalierten Zeitreihenvergleichs** ist auch dem nebenstehenden Beispiel zu entnehmen: Selbst für die sehr lange haltbaren Waren eines Einzelhandels mit Elektroartikeln lässt sich die Summenlogik des monatlichen Wareneinkaufs zum Absatz an den beiden korrelierenden Grafen einwandfrei ableiten. Demgegenüber können die außerordentlichen Schwankungen des Aufschlagsatzes¹⁸⁸ leicht zur Fehldeutung der betriebswirtschaftlichen Unschlüssigkeit verleiten. Abermals bestätigt der Korrelationskoeffizient mit fast +1 den plausiblen Eindruck der zentralen SRP-Zeitreihe und unterstreicht den eklatanten Unterschied der beiden Zeitreihenvarianten.



¹⁸⁶ Insofern muss Kritik an der SRP in der Literatur, die sich auf Aufschlagsatzzeitreihen bezieht (vgl. z.B.: *Bleschick*, „Der kalkulierte Beanstandungsanlass: Kein Nachweis von Mehrergebnissen durch die Summarische Risikoprüfung – Darstellung der Summarischen Risikoprüfung“, DStR 2017, 353: „Zunächst unternimmt er eine sog. betriebswirtschaftliche Auswertung: Hierzu ermittelt er im Wesentlichen mithilfe des Zeitreihenvergleichs einen Rohgewinnaufschlagsatz. Dieser wird auf den gesamten Wareneinkauf des Steuerpflichtigen angewandt. Unter Außerachtlassung von 20 % der höchsten und ggf. von 20 % der niedrigsten Rohgewinnaufschlagsätze ermittelt er den aus seiner Sicht möglichen Umsatz (sog. ‚Quantilschätzung‘). Liegt der für den Außenprüfer mögliche Umsatz oberhalb des erklärten Umsatzes, sieht der Außenprüfer hierin aufgrund betriebswirtschaftlicher Auffälligkeiten einen Klärungsbedarf.“), als gegenstandslos betrachtet werden.

¹⁸⁷ Ebenso: *Wiggen*, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168.

¹⁸⁸ Die hellblauen Balken zum Aufschlagsatz haben keinen direkten Wertbezug zu den Y-Achsen, sondern sind zur Methodengegenüberstellung über einen Faktor an die Relationen der linken Y-Achse angelehnt worden.

Die bisherigen kritischen Aussagen zur eingeschränkten Anwendbarkeit der Zeitreihenanalytik in der Betriebsprüfung¹⁸⁹ können deshalb nicht überzeugen. *Wolenski* hat ebenfalls an einem nachvollziehbaren Beispiel aufgezeigt, dass *geeignete* Zeitreihenformen für diverse Prüfungsfragen sinnvolle Analyseansätze darstellen¹⁹⁰, welche tiefreichende und aussagekräftige Eindrücke zur betriebswirtschaftlichen Logik einer Vielzahl an Besteuerungsdaten liefern, um einerseits Prüffelder betriebsbezogen festzustellen und andererseits dem Gebot der Gesamtbildprüfung¹⁹¹ zu folgen.

Durch die mehrperspektivische Absicherung (s.a. [Pkt. 2.3.](#)) von auffälligen Zeitreihenergebnissen insbesondere unter Verwendung der interaktiven Möglichkeiten, neue Informationen z.B. aus Einlassungen umgehend in die Gesamtbildprüfung einzupflegen (s. [Pkt. 4.2.9.](#)), wird die Aussagegewissheit als „unschlüssig“ beanstandeter Besteuerungsgrundlagen i.S. des § 158 AO abermals verbessert. Darüber hinaus wird in der Literatur der aus Praxiserfahrungen bekannte Zusammenhang auffälliger Zeitreihenergebnisse mit Ordnungsmängeln der Aufzeichnungen benannt¹⁹², wodurch sich eine Indiziengesamtheit¹⁹³ ergibt.

Mit der Quantilsschätzung (s. [Pkt. 4.8.](#)) bietet die SRP-Vorlage BWA zusätzlich eine [Option](#) an, *bei eingetretener Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO* die systematisch erarbeiteten Erkenntnisse zu den Besteuerungsgrundlagen auch für die Festlegung einer möglichst sachgerechten Höhe zu nutzen. Zwar haben Aufschlagsätze innerhalb einer Gesamtbildprüfung mit SRP kaum Bedeutung, im Fall der Verwerfung der Besteuerungsgrundlagen nach § 158 AO werden sie jedoch verwendet, um mit angemessenem Arbeitsaufwand¹⁹⁴ eine **Vollschätzung** durchzuführen.

¹⁸⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 56: „Bei bestimmten Betriebstypen oder in bestimmten betrieblichen Situationen scheidet der Zeitreihenvergleich schon dem Grunde nach als geeignete Schätzungs[- und nach Ansicht des X. Senats wohl auch als Verprobungs-]methode aus.“

¹⁹⁰ *Wolenski*, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, StbG 6/2016, 268, 271.

¹⁹¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

¹⁹² *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25: „In der Praxis wird sich die Frage, ob ein Zeitreihenvergleich allein die Beweisvermutung des § 158 AO widerlegt, [...] dadurch relativieren, dass die [...] aufgedeckten Unregelmäßigkeiten regelmäßig mit Unregelmäßigkeiten in der Kassenbuchführung einhergehen.“

¹⁹³ BGH v. 27.4.2010 – 1 StR 454/09, NSZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

¹⁹⁴ BFH v. 21.2.1990, BFH/NV 1990, 683, Rz. 31.: „Die Geldverkehrsrechnung ist eine arbeitsaufwendige Schätzungsmethode. Die vom Kläger erklärten Buchführungsergebnisse beruhen auf einer formell ordnungswidrigen Buchführung. Daher ist die Anwendung weniger arbeitsaufwendiger Schätzungsmethoden statthaft.“

4.2.1. Betriebsprofil

Die Schlüssigkeitsbeurteilung von Betriebszahlen, -entwicklungen und -abhängigkeiten ist ohne Kenntnis der **wesentlichen Rahmenbedingungen des geprüften Unternehmens** i.d.R. nicht seriös möglich. Für kalkulatorische Gesamtbetrachtungen ist das an Richtsatzspannen von weit über 100 %¹⁹⁵ direkt erkennbar, welche auf die regionalen und einzelbetrieblichen Verhältnisse zurückgehen. Selbst bei den scheinbar klarsten Verprobungsmethoden „Vermögenszuwachs-“ und „Geldverkehrsrechnung“ lassen sich – in der Aufsummierung – wichtige Größen nur anhand einzelfallspezifischer Informationen sachgerecht rekonstruieren wie bspw. der private Lebensstandard.

Obwohl die SRP-Zeitreihenanalytik zentral auf das für Abhängigkeitsfragen besonders geeignete und gegenüber möglichen Störeinflüssen robuste Modell „doppelskaliert größenangepasster Zeitreihenvergleich“ (s. [Pkt. 4.2.](#)) setzt, sind zur richtigen Einordnung der grafischen Auswertungen grundlegende Informationen über den geprüften Betrieb sinnvoll: So können generell z.B. einseitige Preisveränderungen oder punktuell gesamtbildrelevante Einzelsachverhalte wie betriebsbezogen teure Bestandseinkäufe zu Fehlern bei der operativen Periodenzuordnung in Zeitreihenanalysen führen. Deshalb werden zu Beginn einer SRP – neben der Formalprüfung – die wesentlichen Rahmenbedingungen des geprüften Einzelfalls als Betriebsprofil (betriebswirtschaftliches Profil¹⁹⁶) mittels Aktenstudium und Betriebsbesichtigung aufgenommen.

Mit dem Wissen über derartige Sondersachverhalte, die aufgrund ihrer Besonderheit und Bedeutung für den Erfolg den Unternehmensleitenden bekannt sein dürften, lassen sie sich in den SRP-Vorlagen nachvollziehbar interaktiv bereinigen (s. [Pkt. 4.2.9.](#)). Dadurch wird deren Einfluss umgehend ersichtlich, und sie können i.R. der Plausibilitätsbeurteilung berücksichtigt bzw. analytisch auf die betrieblichen Normalverhältnisse korrigiert werden. Ein zusätzlicher Schutz der SRP, durch den das Übersehen gesamtbildrelevanter Sondervorgänge verhindert werden soll, liegt im systematischen Prüfungsduktus „vom groben Überblick zu den Details“ (s. [Pkt. 4.2.2.](#)).

Als Folge des neutral¹⁹⁷ protokollierenden Zusammenstellens gesamtbildrelevanter Informationen aus der betrieblichen Sphäre zu einem **Betriebsprofil** werden von vornherein rea-

¹⁹⁵ S. in der Richtsatzsammlung für das Kalenderjahr 2016 z.B.: Eisdielen mit einem Rohgewinnaufschlagsatz von 233 % bis 488 % bei einem Häufigkeitsschwerpunkt von 335 %; nachlesbar unter: http://www.bundesfinanzministerium.de/Content/DE/Downloads/BMF_Schreiben/Weitere_Steuerthemen/Betriebspruefung/Richtsatzsammlung/001_2.pdf?__blob=publicationFile.

¹⁹⁶ Unter dem Begriff übernommen in: Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 16.

¹⁹⁷ Insofern unzutreffend wiedergegeben von: Bleschick, „Der kalkulierte Beanstandungsanlass: Kein Nachweis von Mehrergebnissen durch die Summarische Risikoprüfung“, DStR 2017, 353, 357: „Im ersten Schritt wird ein sog. betriebswirtschaftliches Profil erstellt. Hierzu identifiziert der Außenprüfer in der Buchführung des

listische (betriebsnahe) Erwartungen an Gesamtbildergebnisse gestellt – im Gegensatz zu generalisierten Branchenerfahrungen mit Übertragungsproblemen. Dabei ergeben sich die zwei **Kernfragen** zur (regelbasierten) Zeitreihenanalytik mit:

- ? Welche grundsätzliche Tendenz in der Entwicklung von Erfolgsgrößen herrscht vor und für welche Betriebsfaktoren ist welche Intensität an Abhängigkeit im Unternehmen anzunehmen: Wareneinkauf zum Absatz, variable Kosten(stellen) zum Umsatz, Kostenveränderungen einzelner Unternehmensteile, Übereinstimmung von vertraglichen Soll- und Istwerten etc.?
- ? Welcher Grad an kaufmännischer Handlungslogik lässt sich aus der Vorbildung und Erfahrung ausführender Personen sowie aus indirekten Merkmalen wie bspw. dem Unternehmensauftritt, der Ablaufplanung oder den Absatzstrategien ableiten?

Im Kontext der – in der Finanzrechtsprechung bisher verbreitetsten – Zeitreihenanwendung „Aufschlagsatzentwicklung zum Warenfluss eines gastronomischen Betriebs“ hat sich auch der BFH zur **zeitlichen Intensität der Abhängigkeit von Betriebsfaktoren** geäußert¹⁹⁸. Grundsätzlich hat eine betriebsbezogen richtige Annahme der Korrelationsintensität wesentliche Bedeutung für die Auswahl einer sachgerechten Vergleichsperiode. Um Fehler in diesem Bereich der Zeitreihenanalytik zu verhindern, arbeitet der systematische Prüfungsduktus der SRP von groben Aggregationsebenen der Daten zu feineren schrittweise hinunter (s. [Pkt. 4.2.2.](#)), so dass widersprüchliche Teilergebnisse zum Ergründen der Ursachen führen.

Die folgenden Aspekte bestimmen die Intensität betrieblicher Abhängigkeiten:

1. Ausgangspunkt des chronologischen Zusammenhangs von betrieblichen Datenreihen ist die **Notwendigkeit der Faktorvorhaltung**: Für den reibungslosen Betriebsablauf und insbesondere zur Erlöserzielung sind die ausreichende Bereitstellung verkäuflicher Ware bzw. Leistung sowie benötigter Kostenpositionen wie z.B. Personal Voraussetzung. Dabei wird die zeitliche Intensität der Korrelation z.B. von der Güte der Einkaufsplanung, der Liquidität, der Haltbarkeit von Waren sowie der La-

Abhängigkeiten und deren Intensität:

- ? Art der Warenbedarfsanalyse (*erfahrungs- vs. erwartungsorientiert*)
- ? Einkaufsrhythmus (*preis- vs. qualitätsorientiert*)
- ? Abrechnungsmodalitäten
- ? Bestandsstrategie und Lagermöglichkeiten
- ? Bildung der Verkaufspreise (*Marktgleichgewicht vs. Produktabgrenzung*)
- ? Substituierbarkeit von Kostenfaktoren
- ? ...

Steuerpflichtigen zwei Faktoren, von denen der eine regelmäßig der Umsatz und der andere der Wareneinkauf ist.“

¹⁹⁸ BFH v. 25.3.2015, BStBl. II 2015, 743, Leitsatz 1: „Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist.“

gerkapazität bestimmt.

Im Fall von variablen Kostenpositionen können vertragliche Vorgaben die Reaktionszeit erheblich verlängern (z.B. Festanstellungen). Zudem bestehen häufig Ersatzmöglichkeiten (bspw. Substituierbarkeit von Festangestellten durch Aushilfen), so dass (zunächst) längere Vergleichsperioden wie Quartale anzuraten sind.

2. Die **Ablauforganisation** eines Betriebs kann Umstände mit sich bringen, die den operativen Zusammenhang zweier Betriebsfaktoren verzerren: Gewohnheitsmäßiges Einkaufsverhalten kann z.B. tatsächlichen Notwendigkeiten tlw. zuwiderlaufen. Auch Abrechnungsmodalitäten wie z.B. die Inrechnungstellung zum Monatsende können bei kurzen Vergleichsperioden wie Wochen relevante Verzerrungen hervorrufen.

Sowohl eine allzu bedarfsferne Einkaufsplanung als auch eine sehr marktunübliche Zahlungsweise riskieren jedoch erfahrungsgemäß die Unternehmensexistenz und müssen deshalb ohne ordnungsgemäße, nachvollziehbare Unterlagen eher als Indikatoren für unschlüssige Besteuerungsaufzeichnungen angesehen werden.

Punktuelle Sondereinflüsse können – nach Bestätigung durch entsprechende Dokumente wie Lieferscheine – interaktiv (s. [Pkt. 4.2.9.](#)) auf Normalverhältnisse korrigiert werden, um die tatsächlichen Auswirkungen festzustellen.

3. Weiterhin kann die **Erfassung der Geschäftsvorfälle** ursächlich für gesamtbildrelevante Störungen sein.

Nach den umfangreichen praktischen Anwendungserfahrungen zur SRP beeinträchtigen geringfügige Ordnungsmängel wie z.B. die Buchung von Wareneinkäufen nach Bezahlung bei üblichen Zahlungszielen das robuste Zeitreihenmodell „doppelskaliert größenangepasster Zeitreihenvergleich“ nicht, was sich an den Extrembeispielen „Aufzeichnung nach Bezahlung“ und „lange haltbare Waren“ unter [Pkt. 4.2.](#) gut nachvollziehen lässt. Dennoch bietet die SRP-Vorlage BWA für den Sonderfall der Einnahmen-Überschuss-Rechnung nach § 4 Abs. 3 EStG die interaktive Option der Vor- bzw. Rückdatierung der ersten Datenreihe um bis zu ±20 Tage (s. [Pkt. 4.2.9.3.](#)). Dadurch lässt sich der tatsächliche Einfluss problemlos und zügig simulieren.

Im Verhältnis einer Schlüssigkeitsprüfung zu gravierenden Ordnungswidrigkeiten verstärken sich die beiden unabhängigen Voraussetzungen der „Beweiskraft der Buchführung (und Aufzeichnungen)“ nach § 158 AO gegenseitig i.S. einer Indiziengesamtheit¹⁹⁹. Entsprechend schwächt auch die „Drei-Stufen-Theorie“ des BFH zur Schätzung die

¹⁹⁹ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweismwürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

Sorgfaltsanforderungen bei Besteuerungsaufzeichnungen, die formell und materiell unrichtig sind, erheblich ab.²⁰⁰ Zudem beschreibt die AO-Kommentierung den aus der Prüfungspraxis hinlänglich bekannten Zusammenhang aus Ordnungswidrigkeit und sachlicher Unrichtigkeit.²⁰¹ Bei erkennbarer (substantieller) Mitwirkungsbereitschaft ermöglichen die interaktiven Funktionen der SRP-Vorlage BWA (s. [Pkt. 4.2.9.](#)) trotzdem in vielen Fällen die gemeinsame²⁰² Herstellung einer ausreichenden Transparenz für die Plausibilitätsbeurteilung nach § 158 AO.

Entsprechend der generell begrenzten Beweismöglichkeiten von Gesamtbildprüfungen stellt § 158 AO nur indizielle Anforderungen an eine sachliche Beanstandung (s. [Pkt. 4.6.](#)), und schlechtes Wirtschaften führt in gegebener Marktkonkurrenz *erfahrungsgemäß* zum Scheitern eines Unternehmens. Dennoch helfen betriebspezifische Kenntnisse über die vorhandene **unternehmerische Kompetenz** bei der Beurteilung von aufgefundenen Regelstörungen in Zeitreihenanalysen. Um fehlendes Unternehmernesschick von dem Verdacht falscher Besteuerungsgrundlagen unterscheiden zu können, stellt die SRP-Vorlage BWA viele interaktive Funktionen zur Verfügung (s. [Pkt. 4.2.9.](#)), mit deren Hilfe Verprobungsauffälligkeiten (gemeinsam) ergründet und – bei nachvollziehbarer, belegter Erklärung – für die weitere Gesamtbildprüfung auf Normalverhältnisse korrigiert werden können.

Die unternehmerische Kompetenz im betreffenden Einzelfall lässt sich anhand des Aktenstudiums, der Betriebsbesichtigung und der Besprechung einer beispielhaften regelbasierten Auswertung einschätzen:

a. **Vorbildung und Erfahrung** handelnder Personen ergeben sich z.T. aus den Steuerakten und werden häufig im Eröffnungsgespräch unaufgefordert berichtet. Dabei ist zu bedenken, dass viele Ausbildungen auch kaufmännische Grundlagen vermitteln.

Praxiserfahrungen aus dem langjährigen Führen eines Unternehmens bzw. das Aufwachsen in einem Familienbetrieb dürften die wirtschaftliche Kompetenz i.d.R. mindestens genauso geschult haben wie ein Wirtschaftsstudium.

Wirtschaftlichkeit des Handelns:

- ? Vorbildung im Betrieb
(Unternehmer oder leitende Angestellte)
- ? Erfahrung im Betrieb
(Unternehmer oder leitende Angestellte)
- ? Marktpräsenz
(Bekanntheitsgrad,
Corporate Identity etc.)
- ? Marketingstrategien
(Absatzwege, innovative
Produkte und Werbung etc.)
- ? ...

²⁰⁰ BFH. v. 25.3.2015, BStBl. II 2015, 743, Rz. 62 ff.

²⁰¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist.“

²⁰² Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

- b. Noch deutlicher lässt sich die unternehmerische Kompetenz aus **indirekten Beobachtungen** ableiten. So sprechen die folgenden Beispiele für ein über den einzelnen Aspekt hinausgehendes Unternehmergeschick mit dementsprechend (weitgehend) schlüssigen Betriebsdaten und -entwicklungen: die erfolgreiche Vermarktung der eigenen Firma bspw. durch einen einheitlichen Marktauftritt mit hohem Wiedererkennungswert (Corporate Identity), eine strukturierte Personal- oder Ablaufplanung, die gelungene Produktabgrenzung vom Massenmarkt mit hohem Preisdruck, die Eröffnung alternativer Absatzwege, kreative Werbung etc.

Da Außenprüfungen – abgesehen von Nachschauen – regelmäßig kaufmännisch lange Zeiträume betrachten, ist eine regelbasierte Gesamtbildprüfung frei von jeder Auffälligkeit eher unrealistisch. In fast jedem Betrieb kommt es irgendwann zu unvorhersehbaren negativen Markteinflüssen, die sich in den Unternehmenszahlen und -entwicklungen als Unschlüssigkeiten wiederfinden lassen. Umso wichtiger für die gesamte Plausibilitätsbeurteilung gem. § 158 AO sind die **Bereitschaft zur und Qualität der Mitwirkung bei der Sachaufklärung**²⁰³.

Sollten vorgefundene Regelstörungen der zentralen SRP-Zeitreihe (doppelskaliert größenangepasster Zeitreihenvergleich, s. [Pkt. 4.2](#)) bei Vorstellung im Unternehmen (offenbar) erstmalig wahrgenommen werden, sind Zweifel an der Richtigkeit der Besteuerungsdaten durchaus begründet, denn das robuste Zeitreihenmodell arbeitet auf Quartals- und Monatsbasis ausschließlich betriebsbezogen existentielle Vorfälle heraus. Demgegenüber kann anhand nachvollziehbarer Erklärungen und quantifizierender Unterlagen die Ursächlichkeit der Faktoren interaktiv untersucht werden (s. [Pkt. 4.2.9](#)). Außerdem sind nach einschneidenden Betriebsproblemen Kompensationsmaßnahmen zu erwarten wie bspw. Absatzaktionen, Kosteneinsparungen etc. Insofern liefert die gemeinsame Auseinandersetzung mit entdeckten Unschlüssigkeiten i.S. der gemeinsamen Sachaufklärungsverantwortung²⁰⁴ aussagekräftige Hinweise auf die Vertrauenswürdigkeit der Besteuerungsgrundlagen insgesamt.

²⁰³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

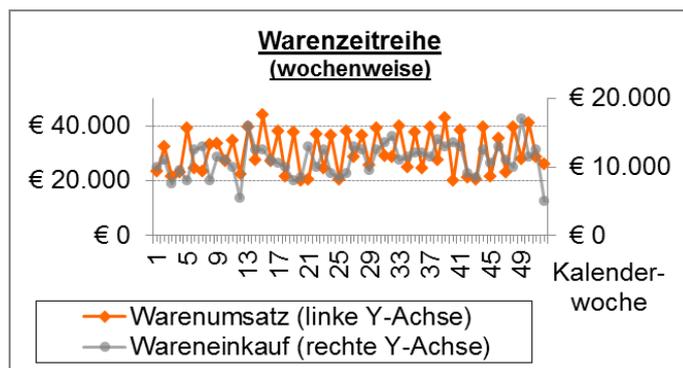
²⁰⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

4.2.2. Prüfungssystematik der SRP-Vorlage BWA

Im Gegensatz zu willkürlichen Methodenzusammenstellungen verfolgt die systematische Verknüpfung geeigneter Zeitreihenanalysen in der SRP-Vorlage BWA²⁰⁵ den Prüfungsduktus „**vom groben Überblick zu den Details**“, um die grundsätzlichen Entwicklungen, Regeln und Muster der Besteuerungsdaten zu verstehen. Prüffelder können so zügig und effektiv aufgedeckt bzw. die Gesamtplausibilität qualifiziert bewertet werden. Die Unterscheidung zwischen systemischen²⁰⁶ Fehlern in Besteuerungsdaten und einzelbetrieblichen Besonderheiten wird verbessert, und es lassen sich punktuelle Prüffelder weitgehend eingrenzen, damit sie rechtlich geprüft und bei Bedarf interaktiv (s. [Pkt. 4.2.9.1.](#)) auf Normalverhältnisse korrigiert werden können. Weitere Datenaufbereitungen der SRP-Vorlage BWA wie z.B. die jahresweise Gegenüberstellung von Monatssummen und Durchschnittswerten der Wochentage dienen der erweiterten mehrperspektivischen Gesamtfallbetrachtung, um die Differenzierung zwischen betrieblichen Sondersituationen und sachlich falschen Besteuerungsgrundlagen sicherzustellen, soweit es die Datenqualität im Einzelfall zulässt.

Der SRP-Zeitreihensystematik liegt das Methodenwissen zu Grunde, dass Zeitreihenvarianten und -ausgestaltungen unterschiedliche Informationen aus den untersuchten Daten herausarbeiten (s. [Pkt. 4.2.](#)). Dabei haben die **Fuzzy-Theorien** (s.a. [Pkt. 2.](#)) eine wichtige Bedeutung, wonach der Grad der Aufsummierung einen wesentlichen Einfluss darauf hat, welche Inhalte erkennbar sind. Diesem Zeitreihenbeispiel eines Gemüsehändlers lassen sich die abweichenden Schwerpunkte der Verbildlichung mit ihren Konsequenzen für die Plausibilitätsbewertung klar entnehmen:

- © Im ersten doppelskalierten Zeitreihenvergleich auf Wochenbasis wird die Erwartung einer engen Abhängigkeit zwischen Wareneinkauf und -umsatz aufgrund der geringen Haltbarkeit des Gemüses nicht bestätigt. Ob der Warennachkauf den Marktbedingungen folgt, ist wegen der **regelmäßigen Schwankungen des Absatzes ca. alle zwei Wochen** nicht nachvollziehbar.

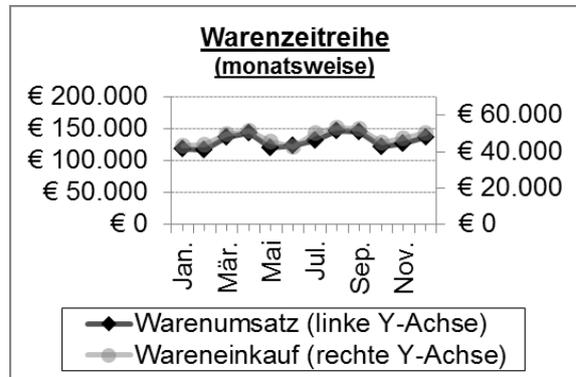


²⁰⁵ Der regelbasierte Prüfungsduktus der SRP-Vorlage BWA wird durch die Abfolge der Arbeitsblätter mit ihren jeweiligen Auswertungen und Funktionalitäten sowie die Kurzanleitung vorgegeben. Die Vorlagenbereiche zur Gesamtbildprüfung sind von denen für die Durchführung einer Schätzung – **nach eingetretener Amtspflicht zur Schätzung gem. § 162 Abs. 2 S. 2 AO** – farblich abgesetzt.

²⁰⁶ Systemische Fehler bzw. Besonderheiten der Datenerfassung und -verarbeitung wie bspw. die verzögerte Aufzeichnung der Einnahmen-Überschuss-Rechnung (§ 4 Abs. 3 EStG) oder falsche Schlüsselungen (z.B. USt-Satz) lassen sich häufig nur oder besser auf stark verdichteten Ebenen wie Quartalen erkennen.

Nach der SRP-Zeitreihensystematik hätte die Untersuchung auf höher verdichteten Summenebenen anfangen sollen und wäre erst schrittweise zu den sehr detailreichen, leicht überfordernden Analysen kurzer Vergleichsperioden fortgesetzt worden.

- © Bereits auf der Monatsebene des zweiten doppelskalierten Zeitreihenvergleichs ist von den starken Schwankungen des Warenumsatzes nichts mehr zu sehen, wodurch die nahezu perfekte Korrelation beider Datenreihen in den Vordergrund tritt. Der offensichtliche Widerspruch bei der Zeitreihenauswertung führt dazu,

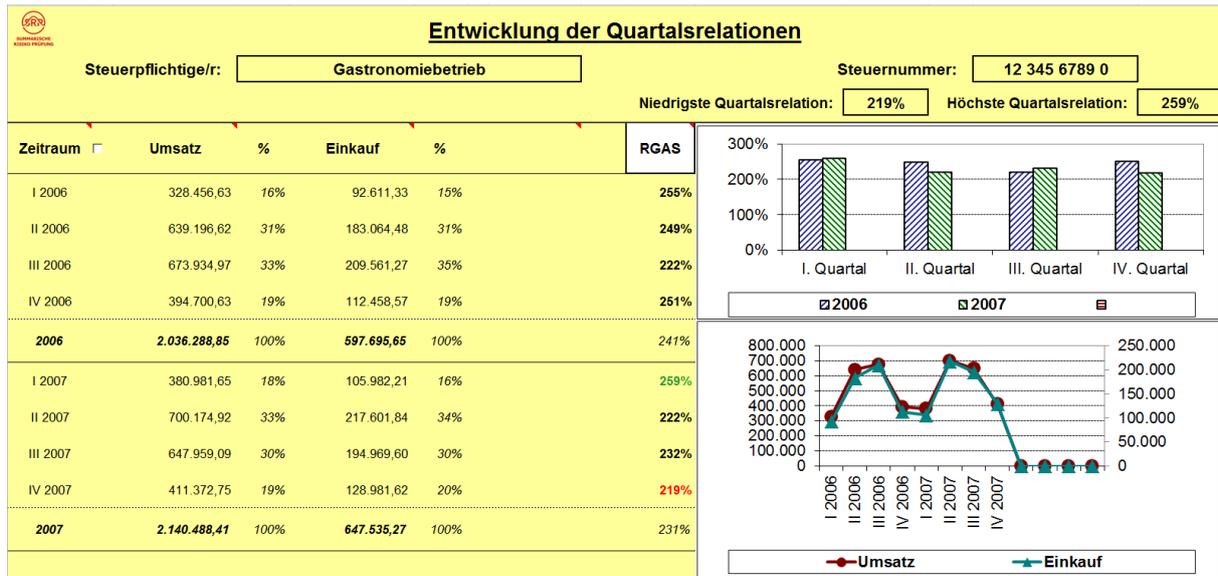


sich die Einzeldaten zum Umsatz chronologisch sortiert mit dem Fokus auf besondere Ereignisse alle zwei Wochen anzuschauen. Mit diesem Vorwissen fällt sofort auf, dass im geprüften Unternehmen jeweils zu Monatsmitte und -ende Rechnungen an bestimmte Kunden gestellt werden. Somit erklärt ein eher untypischer Abrechnungsmodus für die betreffende Branche, dass eine wöchentliche Zeitreihenanalyse ohne periodengerechte Aufteilung der Rechnungsumsätze nicht für die Schlüssigkeitsbewertung in Betracht kommt. Die geeignete monatliche Warenflusszeitreihe legt mit ihrer sehr guten Handlungslogik keine weitere Intensivierung nahe.

Im Ergebnis wird durch das systematische Vertiefen der SRP-Zeitreihen vom groben Überblick langer Vergleichsperioden zu den detaillierteren Perspektiven kürzerer nicht nur eine höhere Effektivität bei der Aufdeckung von betriebsspezifischen Prüffeldern erreicht, sondern eine Plausibilitätsbeurteilung gem. § 158 AO gegen falsche Auswertungsinterpretationen abgesichert – auch im Hinblick darauf, dass protokollierte Betriebsprofile (s. [Pkt. 4.2.1.](#)) unvollständig sein können.

Neben dem schrittweisen Arbeiten mit der Vergleichsperiode werden im SRP-Zeitreihensystem unterschiedliche Zeitreihenvarianten mit ihren jeweiligen Stärken für eine effiziente, weitreichende **Eingrenzung aufgefundener Regelstörungen zu ursächlichen Sachverhalten** genutzt. Im Anschluss können Einzelprüffelder auf die richtige rechtliche Behandlung untersucht werden. Zudem lassen sie sich bei Bedarf interaktiv (s. [Pkt. 4.2.9.1.](#)) auf Normalverhältnisse korrigieren, sofern sie eine vorzunehmende Schlüssigkeitsbewertung nach § 158 AO stören.

Am Beispiel des Warenflusses eines größeren gastronomischen Betriebs soll die SRP-Zeitreihensystematik vom Verstehen der Größenzusammenhänge über das Erkennen bis zum Eingrenzen von konkreten Prüffeldern beschrieben werden:



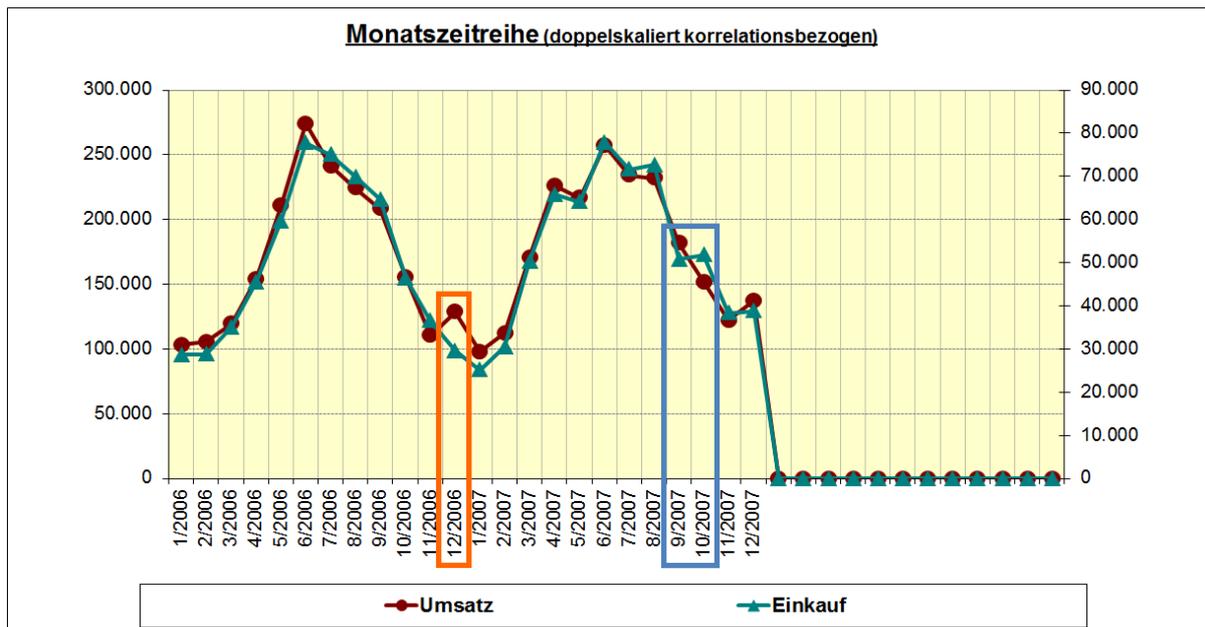
Für die grobe Verbildlichung der quartalsweisen Entwicklungen und des Zusammenhangs beider Datenreihen werden zwei Ansichten zur Verfügung gestellt:

1. Die jahresweise Gegenüberstellung des Verhältnisses aus den Quartalssummen soll insbesondere jahreszeitliche, wiederkehrende Strukturveränderungen anzeigen. Im Beispielsfall ist keine gleichmäßige Erhöhung oder Absenkung des Aufschlagsatzes zu bestimmten Jahreszeiten erkennbar, woraus sich schließen lässt, dass die veränderten Nachfragegewohnheiten der Kunden über das Jahr keinen relevanten Einfluss auf visualisierte Beurteilungen der Warenflussplausibilität haben.
2. Mit dem – demgegenüber streng chronologischen – doppelskaliert größenangepassten Zeitreihenvergleich wird die Reaktion des Warennachkaufs im Betrieb auf die Absatzverhältnisse überschlägig²⁰⁷ wiedergegeben. Im Beispielsunternehmen zeigt sich eine (nahezu) perfekte Anpassung an die stark veränderte Nachfragesituation der einzelnen Jahreszeiten.

Im nächsten Schritt wird mit dem doppelskaliert größenangepassten Zeitreihenvergleich, der für die Beurteilung der Korrelation (Abhängigkeit) zweier Datenreihen besser geeignet ist als andere Zeitreihenmodelle (s. [Pkt. 4.2.](#)), die Prüfungsmaßnahme vertieft, indem mit dem Monat ein deutlich kürzeres Vergleichsintervall gewählt wird. Weiterhin ist der enge Zusammenhang beider Betriebsfaktoren unproblematisch abzulesen. Als Folge der wesentlich detaillierteren Analyse der Besteuerungsdaten sind nun aber auch eine klare **Ablösung des Wareneinkaufs 12/2006 vom ansteigenden Umsatz** sowie eine **Zuordnungsungenauigkeit**²⁰⁸ in den Monaten 9-10/2007 zu beobachten:

²⁰⁷ Infolge des statistischen Ausgleichs von einzelnen Ausreißerwerten durch die Aufsummierung über lange Quartalszeiträume ist nur eine sehr grobe Beurteilung der Handlungslogik möglich. Einzelprüffelder sind auf dieser Ebene i.d.R. nicht sichtbar.

²⁰⁸ In Rechtsprechung (vgl. z.B.: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 47: „Dies zeigt, dass bereits eine Veränderung in der Zuordnung der Einkäufe nur einer einzigen Woche erhebliche Auswirkungen auf das Ge-



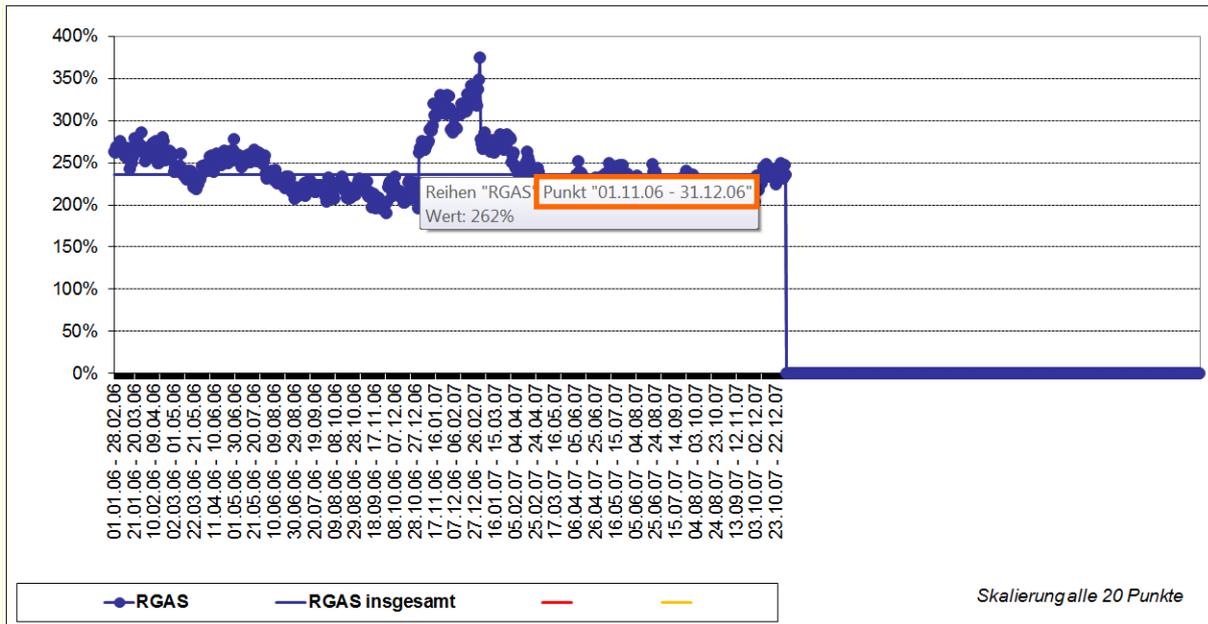
Während sich der (relativ) überhöhte Wareneinkauf in 10/2007 zumindest tlw. durch den Vormonat aufhebt, offenbart sich mit dem fehlenden Anstieg in 12/2006 eine Regelstörung, die als Einzelprüffeld nachverfolgt wird.

Da Wochenzeitreihen in den meisten Branchen zu störanfällig für alltägliche Betriebsabläufe wie z.B. Bestandseinkäufe sind, nutzt die SRP zur weiteren Konkretisierung der aufgedeckten Auffälligkeit ein gleitendes²⁰⁹ Zeitreihenmodell. Hierbei wird die Entwicklung der Verhältnisgröße beider Datenreihen zwar für einen robust langen Zeitraum wiedergegeben, jedoch immer nur um einen Tag weitergetragen. In der Standardeinstellung der SRP-Vorlage BWA „2 Monate“ (s. [Pkt. 4.2.6.](#)) folgt dann auf den Aufschlagsatz aus Wareneinkäufen und -umsatz zum Startintervall „1.1. bis 28.2.“ derjenige für „2.1. bis 1.3.“ und danach für „3.1. bis 2.3.“ etc. Diese Auswertungsform bewirkt, dass ein besonderer Einzelsachverhalt, der ursächlich für die beobachtete Regelstörung ist, an zwei Punkten des Diagramms den Grafen abrupt verzerren muss: beim Ein- und Austreten aus der Vergleichsperiode.

Im Beispielfall ist dementsprechend ein plötzliches Ansteigen des Aufschlagsatzes mit dem **hervorgehobenen Intervall „1.11. bis 31.12.2006“** zu sehen:

samtergebnis der Zehn-Wochen-Periode hat.“) und Literatur (vgl. z.B.: *Wiggen*, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, 171: „Von großer Bedeutung für die Qualität und Beweiskraft des ZRV ist, dass der Schritt zum zutreffenden ‚Periodenwareneinsatz‘ gelingt.“) wird vor Zuordnungsunschärfen in Zeitreihenanalysen vehement gewarnt, obwohl die verbildlichte Schlüssigkeitsprüfung gerade nicht in der Beurteilung einer einzelnen Zahl besteht, sondern der chronologische Gesamteindruck vor solchen Fehlschlüssen wie im Beispiel schützt.

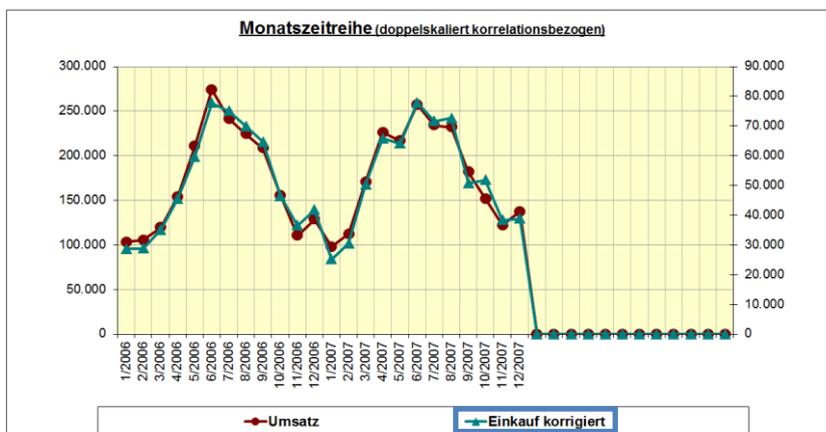
²⁰⁹ Auch „rollierende Zeitreihe“, vgl.: *Harle/Olles*, „Die moderne Betriebsprüfung“, NWB Verlag, 3. Auflage 2017, Rz. 1366 ff.



Nach diversen Punkten auf einem höheren Niveau fällt der Graf dann wieder abrupt ab. Dieses besondere Verhalten der Gleitschlittenzreihe erklärt sich dadurch, dass ein außergewöhnlich niedriger Wareneinkauf vom 31.12.2006 erstmalig die betreffende Vergleichsperiode beeinflusst. Da die Folgeintervalle aber immer nur einen Tag weitergleiten, bleibt der kritische Tag etliche Punkte lang in den jeweiligen Aufschlagsätzen wirksam, bis sein Einfluss mit dem Intervall „1.1. bis 28.2.2007“ plötzlich verschwindet.

Durch eine gezielte Filterung im Wareneinkauf des 31.12.2006 lässt sich zügig ein hoher Negativbetrag auffinden. Nach Belegsichtung wird deutlich, dass es sich um eine Bonuszahlung handelt, die nicht über ein Extrakonto transparent abgegrenzt worden ist. Somit hat die Aufklärung zur Konsequenz, dass die Behandlung etwaiger Boni im Folgejahr intensiv geprüft wird, weil eine vergleichbare Ablösung des Wareneinkaufs vom -umsatz dort im doppelskalierten Monatszeitreihenvergleich nicht vorkommt.

Die Gegenkontrolle, ob der identifizierte Sondersachverhalt vollumfänglich die beobachtete Zeitreihenauffälligkeit verursacht hat, mit



interaktiv (s. [Pkt. 4.2.9.1.](#)) egalisiertem Warenbonus vom 31.12.2006 führt nun zum schlüssigen Gleichlauf des Warennachkaufs zur aufoktroierten Marktsituation „Nachfrage“.

Insgesamt ermöglicht es die systematische SRP-Zeitreihenanalyse, entlang des Duktus „vom groben Überblick zu den Details“ und unter planvoller Kombination unterschiedlicher

Zeitreihenvarianten, in vielen tausend Einzelvorgängen die betriebsspezifischen Prüffelder zielgenau und zügig herauszuarbeiten. Auch das Ergebnis einer Plausibilitätsbeurteilung wird dadurch wesentlich qualifizierter, weil – neben der besseren Eignung des zentralen SRP-Modells „doppelskaliert größenangepasster Zeitreihenvergleich“ (s. [Pkt. 4.2.](#)) – die im Zusammenhang mit Zeitreihenanalysen stets problematisierten Bestandseinkäufe wirksam identifiziert und umverteilt werden können.

4.2.3. Jahressummen

Aus der jahresweisen Aufsummierung von Besteuerungsdaten können i.d.R. keine neuen Erkenntnisse gegenüber dem Veranlagungsverfahren gewonnen werden. Deshalb liegen die **wesentlichen Funktionen** des Tabellenblatts „Jahreszahlen“ darin, die hineinkopierten Daten zügig auf Vollständigkeit zu kontrollieren (s.a. [Pkt. 4.](#)) und die evtl. zeitreihenrelevanten Informationen „Bestandsveränderungen“ und „Entnahmen“ zu ergänzen:

Jahreszahlen				
Steuerpflichtige/r:	Café		Steuernummer:	12 345 6789 0
	2007	2008	2009	
oberer Richtsatzwert:	400%	400%	400%	
mittlerer Richtsatzwert:	257%	257%	257%	
unterer Richtsatzwert:	186%	186%	186%	
Umsatz:	72.273,01	73.121,67	81.548,09	
Einkauf:	31.545,43	33.551,15	31.563,40	
Bestandsveränderung:	0,00	0,00	0,00	
Entnahmen:	0,00	2.895,00	3.108,00	
Einsatz:	31.545,43	30.656,15	28.455,40	
RGAS:	129%	139%	187%	

Die jährlichen Warenentnahmen können über **ein Anhakfeld** auf dem Tabellenblatt „Monatszeitreihen“ für die quartals- und monatsweisen Zeitreihenanaly-

5	2007		Wenn die Entnahmen vom Tabellenblatt „Jahreszahlen“ berücksichtigt werden sollen, bitte hier einen Haken setzen.			
6	Monat	Umsatz	Einkauf			
7	Jan	5.717,99	8%	2.728,08		
8	Feb	4.534,49	6%	2.124,11	7%	113%
9	Mär	6.036,05	8%	2.320,54	7%	160%
10	Apr	5.586,92	8%	2.646,79	8%	111%
						6.568,53
						5.963,08
						5.544,09

sen gleichmäßig auf die (vorhandenen) Perioden verteilt werden.²¹⁰ Eine vergleichbare Option besteht für die Bestandsveränderungen nicht, weil diese Art Auf- und Abbau des Warenbestands nicht dem üblichen Geschäftsablauf entspräche. Über die interaktiven Korrekturoptionen auf dem Tabellenblatt „Vergleichsdaten“ lassen sie sich aber – unter Mitwirkung des Stpfl. – problemlos und schnell auf die tatsächlichen Verhältnisse abändern (s. [Pkt. 4.2.9.1.](#)).

Obwohl der externe Vergleich als größtenteils²¹¹ Verprobungsmethode in der SRP keine entscheidende Bedeutung hat, lassen sich die Richtsatzwerte²¹² für den Aufschlagsatz zur Voll-

²¹⁰ Ist diese Verteilung der Jahresentnahmen vor Aktivierung der Quantilsschätzung (s. [Pkt. 4.8.](#)) nicht erfolgt, wird auf dem Tabellenblatt „Schätzung“ formatseitig darauf hingewiesen.

²¹¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 17: „Als größtenteils Methode zur Beanstandung der sachlichen Richtigkeit des Buchführungsergebnisses erweist sich der äußere Betriebsvergleich [...] Ver-

ständigkeit des Eindrucks jahresweise eintragen. Wird dieser Rahmen wie im oberhalb abgebildeten Beispiel unterschritten, weist die rote Feldfarbe darauf hin.²¹³

4.2.4. Quartalsrelationen

Gem. dem generellen Zweck der Statistik, aus einer größeren Datenmenge die Kerninformationen herauszufiltern²¹⁴, unterdrücken quartalsweise Aufsummierungen einer Vielzahl von betrieblichen Vorgängen enthaltene Besonderheiten zu Gunsten grundsätzlicher Größenordnungen, Tendenzen, Regeln und Muster. Diese grobe „Hubschrauberperspektive“ ist zu Beginn einer Prüfung als **Erstorientierung** wichtig, weil durch das Wissen über die grundsätzlichen Eigenschaften von Betriebsdaten die Unterscheidung zwischen systemischen²¹⁵ Unternehmensbesonderheiten und -fehlern ermöglicht wird. Andererseits darf von einer regelbasierten Gesamtbildprüfung bspw. zum Warenfluss alleine auf Quartalsebene keine ausreichende Gründlichkeit erwartet werden, denn die meisten betriebsbezogenen Einzelprüffelder werden unter der hochgradigen Verdichtung unsichtbar.

Die SRP verwendet zur Groborientierung **zwei zielgerichtete Ansichten**:

1. Das Zusammenfassen beider Datenreihen zu einem Verhältnisfaktor²¹⁶ mit jahresweiser Gegenüberstellung der Quartalswerte vermittelt einen schnellen Eindruck, ob unterjährige Strukturveränderungen gesamtbildrelevant sein können. So würde bspw. bei einer Kostenzeitreihe mit jahreszeitlichem Austausch wesentlicher, unterschiedlich teurer Kostenpositionen das Verhältnis wiederkehrend – quasi „wellenartig“ – schwanken. Trotz der Stabilität des doppelskaliert größenangepassten Zeitreihenvergleichs (s. [Pkt. 4.2.](#)) ist die Kenntnis starker struktureller Saisoneffekte, welche sogar auf robuster Quartalsebe-

gleichen mit anderen Betrieben kommt regelmäßig ein geringerer Vergleichswert als inneren Betriebsvergleichen [...] zu“.

²¹² Die – für den Dreijahreszeitraum „2007 bis 2009“ unveränderten – Richtsätze für Cafés können nachgelesen werden unter:

https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Steuern/Weitere_Steuerthemen/Betriebspruefung/Richtsatzsammlung_Pauschbetrage/richtsatzsammlung-pauschbetragee-fuer-unentgeltliche-wertabgaben-anlage057.html.

²¹³ Eine entsprechende bedingte Formatierung ist auf dem Tabellenblatt „Quartalsrelationen“ und bei den Zeitreihenvergleichen des Tabellenblatts „Monatszeitreihen“ nicht hinterlegt, weil die Plausibilitätsprüfung der SRP nicht an der Konstanz des (empfindlichen) Aufschlagsatzes ansetzt, sondern an der wesentlich robusteren Handlungslogik einer angemessenen Reaktion veränderbarer Betriebsfaktoren auf Markteinflüsse wie bspw. der Anpassung des Wareneinkaufs auf die Nachfrage (s. den Abschnitt „Methodenkenntnis“ unter Pkt. 4.2.).

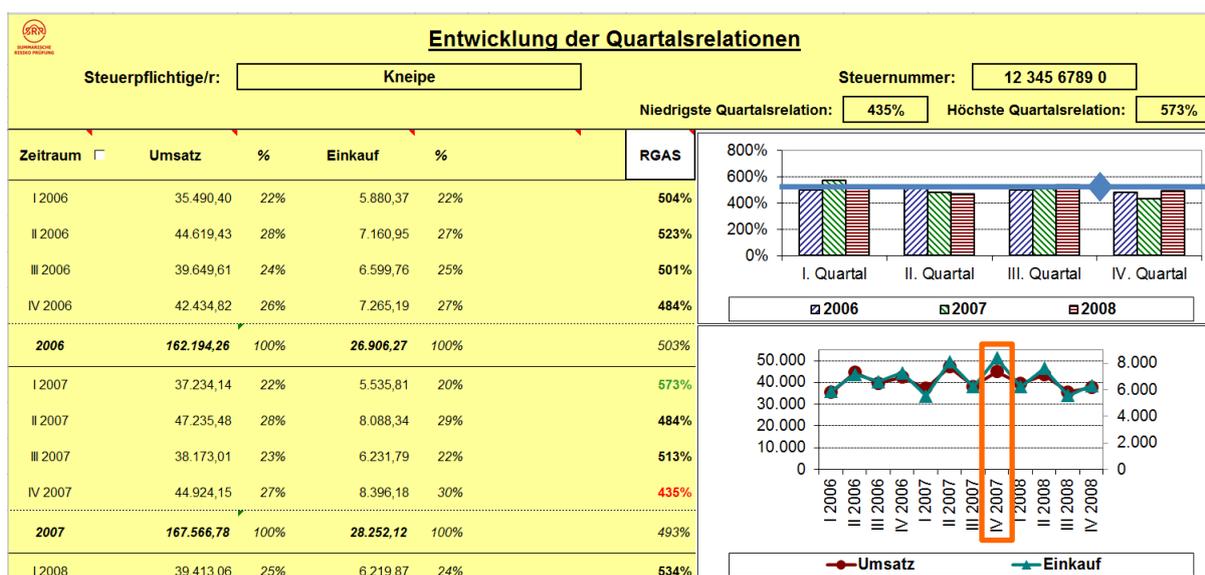
²¹⁴ Auf diesem Zweck basiert auch die betriebswirtschaftliche Beratung und Planung mit der quartalsweisen und monatlichen Zusammenfassung umfangreicher entscheidungsrelevanter Daten.

²¹⁵ Systemische Fehler bzw. Besonderheiten der Datenerfassung und -verarbeitung wie bspw. die verzögerte Aufzeichnung der Einnahmen-Überschuss-Rechnung (§ 4 Abs. 3 EStG) oder falsche Schlüsselungen (z.B. USt-Satz) lassen sich häufig nur oder besser auf stark verdichteten Ebenen wie Quartalen erkennen.

²¹⁶ Der Verhältnisfaktor wird i.S. eines Aufschlagsatzes ermittelt mit: $\frac{\text{Datenreihe 2} - \text{Datenreihe 1}}{\text{Datenreihe 1}}$ und prozentual ausgegeben.

ne offensichtlich sind, für die sachgerechte Deutung insbesondere von Monatszeitreihen wichtig.

- Zusätzlich werden beide Datenreihen mit dem doppelskalierten Zeitreihenvergleich gegenübergestellt, weil dieser am besten die Reaktion der steuerbaren Größe auf die andere – also z.B. des Wareneinkaufs auf den Absatz – wiedergibt (s. den Vergleich von drei Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)). Damit wird bereits auf der hochverdichteten Quartalsebene der Fokus auf die sinnvollste Plausibilitätsfrage i.S. des § 158 AO gelegt: „Hat die Unternehmensleitung so erfolgsorientiert auf Markteinflüsse reagiert, dass das Gesamtbild der Besteuerungsdaten keinen Anlass bietet, an ihrer sachlichen Richtigkeit zu zweifeln?“



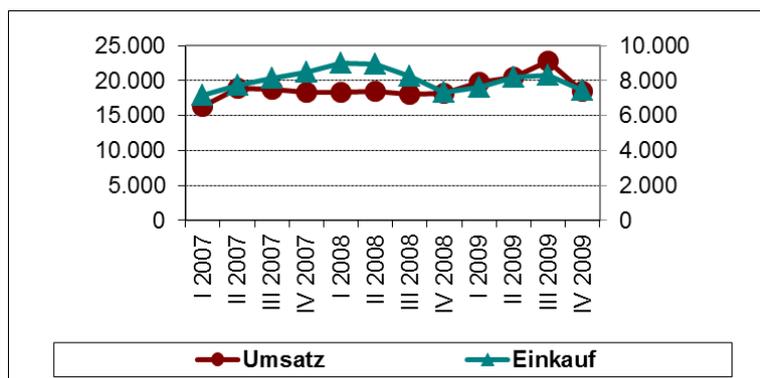
Für den Warenfluss der exemplarisch abgebildeten Kneipe sind der oberen Grafik keine jahreszeitlichen „Wellenbewegungen“ des Aufschlagsatzes zu entnehmen, die auf unterjährig erheblich veränderte Getränkeassortimente mit abweichenden Margen hinweisen würden. Die Summenbetrachtung vieler Warenumsätze und -einkäufe pro Quartal, infolge der die Besonderheiten größtenteils „abgeschliffen“ werden, ergibt **Aufschlagsätze mit einer nur geringen Schwankungsbreite**.

Dem streng chronologischen Zeitreihenvergleich der zweiten Abb. nach ergeben sich jedes Jahr saisonartige Zu- und Abnahmen des Warendurchflusses. Dabei folgt der Getränke- nachkauf der Nachfrage weitgehend in Richtung und Intensität. Auch die verzögerte Aufzeichnung des Wareneinkaufs aufgrund der vorliegenden Einnahmen-Überschuss-Rechnung gem. § 4 Abs. 3 EStG hat den Zusammenhang der Betriebsfaktoren bei dem robusten Zeitreihenmodell nicht stören können. Dadurch tritt die **Einzelablösung IV/2007** hervor, wo der (relativ) überhöhte Einkauf auch im Folgequartal offenbar nicht verbraucht worden ist²¹⁷. Die

²¹⁷ Besonders in der Literatur (vgl. z.B.: *Wiggen*, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, 171: „Von großer Bedeutung für die Qualität und Beweiskraft des ZRV ist, dass der Schritt zum zutreffenden

singuläre Auffälligkeit kann nun nach dem Prüfungsduktus der SRP-Vorlage BWA systematisch nachverfolgt werden, um den ursächlichen Vorgang rechtlich zu würdigen und ggf. für die weitere Gesamtbildprüfung interaktiv (s. [Pkt. 4.2.9.1.](#)) auf Normalverhältnisse zu berichtigen. *Durch die sich anschließende Intensivierung der Zeitreihenanalysen auf die empfindlichere Monatsbasis wird eine annehmbare Gründlichkeit bei der Aufdeckung betriebsspezifischer Prüffelder erreicht.*

Offensichtliche Regelablösungen



zeitreihenvergleichs wie im nebenstehenden Beispiel eines Cafés mit langfristig erheblich verstärktem Wareneinkauf begründen schwerwiegende Zweifel an der inhaltlichen Richtigkeit der Besteuerungsdaten, weil die üblichen Erklärungsansätze²¹⁸

für Zeitreihen auffälligkeiten zur Zerstreung nicht geeignet sind:

- ! Einerseits ist das Modell des doppelskaliert größenangepassten Zeitreihenvergleichs stabil gegenüber Zuordnungsunschärfen, so dass der Zusammenhang bspw. von Warenflussdaten selbst bei Einnahmen-Überschuss-Rechnung gem. § 4 Abs. 3 EStG oder für lange haltbare Waren weiterhin problemlos zu erkennen ist (s. *das Beispiel oben zu diesem Pkt. bzw. die Monatszeitreihe im Abschnitt „Robustheit des doppelskalierten Zeitreihenvergleichs“ unter [Pkt. 4.2.](#)*).
- ! Andererseits sind Quartalssummen sehr große – und damit robuste – Bezugsparameter, wodurch mögliche Zuordnungsfehler „kleingerechnet“ werden.²¹⁹
- ! Die – in der Literatur gefürchteten²²⁰ – hochwertigen Bestandseinkäufe mit verzögertem Verbrauch bewirken punktuelle Ablösungen im Gleichlauf der Grafen „Warenumsatz“ und „Wareneinkauf“. Nach der SRP-Zeitreihensystematik werden derartige Einzelauffälligkeiten schrittweise eingegrenzt und bis zu den verursachenden Vorgängen nachver-

„Periodenwareneinsatz“ gelingt.“) wird vor Zuordnungsunschärfen in Zeitreihenanalysen vehement gewarnt, obwohl die verbildlichte Schlüssigkeitsprüfung gerade nicht in der Beurteilung eines einzelnen Diagrammpunkts besteht, sondern der chronologische Gesamteindruck vor solchen Fehlschlüssen schützt.

²¹⁸ Vgl. z.B.: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 47: „Dies zeigt, dass bereits eine Veränderung in der Zuordnung der Einkäufe nur einer einzigen Woche erhebliche Auswirkungen auf das Gesamtergebnis der Zehn-Wochen-Periode hat.“

²¹⁹ Wolenski, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, Stbg 6/2016, 268, mit einem umfangreichen Rechenbeispiel auf S. 269 zur Bedeutung der Größe der jeweiligen Bezugssummen (Wochen, Monate, 10-Wochen-Intervalle, Quartale etc.) auf die Störproblematik.

²²⁰ Wiggen, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, 171: „Von großer Bedeutung für die Qualität und Beweiskraft des ZRV ist, dass der Schritt zum zutreffenden ‚Periodenwareneinsatz‘ gelingt.“

folgt (s. [Pkt. 4.2.2.](#)). Die aufgefundenen, gesamtbildrelevanten Sachverhalte können nachvollziehbar interaktiv (s. [Pkt. 4.2.9.1.](#)) auf Normalverhältnisse korrigiert werden, so dass sie das Schlüssigkeitsurteil i.S. des § 158 AO nicht beeinflussen.

- ! Darüber hinaus zeigt die ständige Prüfungspraxis, dass offenkundige Implausibilitäten in Bestimmungsaufzeichnungen (nahezu) nie losgelöst von Formalmängeln und anderen materiellen Indizien vorkommen.²²¹ Darum wird die Beurteilung der Aussagekraft einer so auffälligen Zeitreihenauswertung wie im Beispielfall regelmäßig im Kontext von Indizien Gesamtheiten²²² stattfinden. Hierbei können die interaktiven (s. [Pkt. 4.2.9.](#)) Optionen der SRP-Vorlage BWA praktikabel zur Einarbeitung substantiiertes Einlassungen i.R. der gemeinsamen Aufklärungsverantwortung²²³ genutzt werden. Auf die Weise lassen sich reine Schutzbehauptungen von tatsächlichen außergewöhnlichen Begründungen effektiv und zügig unterscheiden.

4.2.5. Monatszeitreihen

Dass **monatliche Summen** zur Bewertung von betrieblichen Entwicklungen und Abhängigkeiten

- ✓ auf der einen Seite eine ausreichende Robustheit – also geringe Störanfälligkeit gegenüber eher alltäglichen Sondersituationen – haben und
- ✓ auf der anderen Seite empfindlich genug zum Aufzeigen wichtiger Veränderungen und Einflüsse sind,

wird aus der seit Jahrzehnten üblichen Praxis in Unternehmensplanung und -beratung ersichtlich, Monats-BWAs wichtigen Entscheidungen zu Grunde zu legen. Werden die monatlich aggregierten Betriebsdaten zusätzlich durch eine zielführende Zeitreihenvariante verbildlicht, können existentielle Schlussfolgerungen z.B. zu Investitionen verantwortungsvoll abgeleitet werden. Folgerichtig muss eine [geeignete](#) (s. [Pkt. 4.2.](#)) Zeitreihenanalyse der Monatssummen zweier betrieblicher Datenreihen grundsätzlich auch für eine Plausibilitätsbetrach-

²²¹ Madauß, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25: „In der Praxis wird sich die Frage, ob ein Zeitreihenvergleich allein die Beweisvermutung des § 158 AO widerlegt, [...] dadurch relativieren, dass die [...] aufgedeckten Unregelmäßigkeiten regelmäßig mit Unregelmäßigkeiten in der Kassenbuchführung einhergehen.“

²²² BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

²²³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

tung i.S. des § 158 AO in Betracht kommen – insbesondere wenn sie gegen ungerechtfertigte Beanstandungen durch die Prüfungssystematik „vom groben Überblick zu den Details“ (s. [Pkt. 4.2.2.](#)) und durch interaktive Optionen (s. [Pkt. 4.2.9.](#)) zur Bereinigung von gesamt-bildrelevanten Sondersituationen abgesichert wird.

Auf dem Tabellenblatt „Monatszeitreihen“ der SRP-Vorlage BWA werden **drei Zeitreihenvarianten** entsprechend ihrer unterschiedlichen Vorzüge für bestimmte Zwecke mit zielgerichtet unterstützenden Funktionalitäten zur Verfügung gestellt:

1. Als zentrales SRP-Zeitreihenmodell gibt der doppelskaliert größenangepasste Zeitreihenvergleich die Abhängigkeit zweier Datenreihen auch unter schwierigen Bedingungen wie der Einnahmen-Überschuss-Rechnung nach § 4 Abs. 3 EStG oder längerer Haltbarkeit der Waren korrekt wieder (s. *die entsprechenden Beispiele unter [Pkt. 4.2.](#)*). Als statistisches Maß der Gleichläufigkeit kann der Korrelationskoeffizient²²⁴ für beliebige zusammenhängende Zeiträume als Gegenkontrolle des visuellen Eindrucks ausgegeben werden. Zudem bietet die Funktion der Konjunkturanpassung die Möglichkeit, einseitige Preiseffekte mit evtl. gesamt-bildrelevanter Bedeutung interaktiv auszugleichen (s. *das nachfolgende Beispiel unter [Pkt. 4.2.5.1.](#)*), um zwischen extremen Markteinflüssen und unschlüssigem Handeln unterscheiden zu können.
2. Der anschließende einzelskalierte Zeitreihenvergleich wird besonders für Übereinstimmungsprüfungen benötigt (s. [Pkt. 5.3.](#)), wenn bspw. Datenschnittstellen auf Fehlerfreiheit untersucht oder Steuerungsdaten anhand externer Datenbanken abgeglichen werden sollen. Dazu werden die monatlichen und jährlichen Differenzen der beiden Datenreihen zusätzlich in einer Tabelle dargestellt. In seltenen Fällen kann dieser Zeitreihentyp für eine Plausibilitätsprüfung ergänzend hilfreich sein, weil die unternehmerische Bedeutung einer Regelablösung deutlicher hervortritt (s. *das nachfolgende Beispiel unter [Pkt. 4.2.5.2.](#)*). Weil durch die sekundäre Y-Achse des doppelskalierten Zeitreihenvergleichs nicht immer automatisch die perfekte Größenanpassung erreicht wird, lässt sich dieser Zeitreihenvergleich auf relative²²⁵ Darstellung inkl. der Funktion umstellen, die erste Datenreihe in der Abbildungshöhe an die zweite anzupassen.
3. **Die abschließende Zeitreihenanalyse der Verhältnisgröße²²⁶ aus beiden Datenreihen wie bspw. der Aufschlagsatz beim Warenfluss hat i.R. der SRP keine entscheidende Prüfungsbedeutung, weil sie erfahrungsgemäß empfindlich gegenüber Störeinflüssen ist** (s.

²²⁴ Als lineares Abhängigkeitsmaß nimmt der Korrelationskoeffizient Größen von „-1“ für die 100 %-ige Gegenläufigkeit der Grafen über „0“ bei nicht vorhandener Beeinflussung bis zu „+1“ für eine vollständig (positive) Gleichläufigkeit ein.

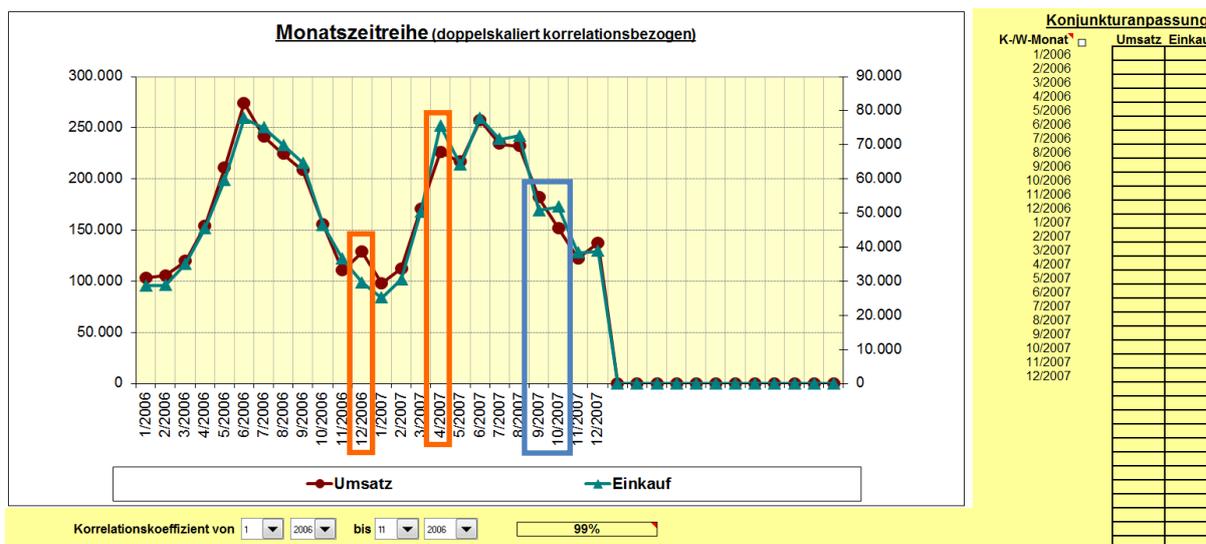
²²⁵ Die Monatswerte werden als Anteile der jeweiligen Jahressumme ausgedrückt.

²²⁶ Der Verhältnisfaktor wird i.S. eines Aufschlagsatzes ermittelt mit: $\frac{(\text{Datenreihe 2} - \text{Datenreihe 1})}{\text{Datenreihe 1}}$ und prozentual ausgegeben.

den Modellvergleich der gängigsten Zeitreihenanalysen im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)). Der Zweck dieses Teils der „Monatszeitreihen“ liegt darin, bei der Durchführung einer Quantilsschätzung (s. [Pkt. 4.8.](#)) – unter vorliegender Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO – durch tabellarisches Ranking der einzelnen Monatswerte und Verbidlichung die Nachvollziehbarkeit bei der Auswahl der Schätzgröße herzustellen. Zudem lassen sich einzelne – insbesondere inoperative – Monate aus der internen Ermittlung der 80 %-Quantile ausschließen.

4.2.5.1. Doppelskaliert größenangepasster Zeitreihenvergleich

Aufgrund seiner **Robustheit gegenüber möglichen Störeinflüssen** (s. die Beispiele zur Einnahmen-Überschuss-Rechnung und zu lange haltbaren Waren unter [Pkt. 4.2.](#)) ist der doppelskaliert größenangepasste Zeitreihenvergleich i.R. der SRP auch auf Monatsbasis das zentrale Mittel zur Plausibilitätsbeurteilung von Besteuerungsgrundlagen. Dazu werden zunächst betriebsspezifische Prüffelder mit ausreichender Gründlichkeit gem. SRP-Zeitreibensystematik (s. [Pkt. 4.2.2.](#)) aufgedeckt, um die Beurteilungsgrundlage für die Schlüssigkeit nachvollziehbar von Störfaktoren zu säubern.



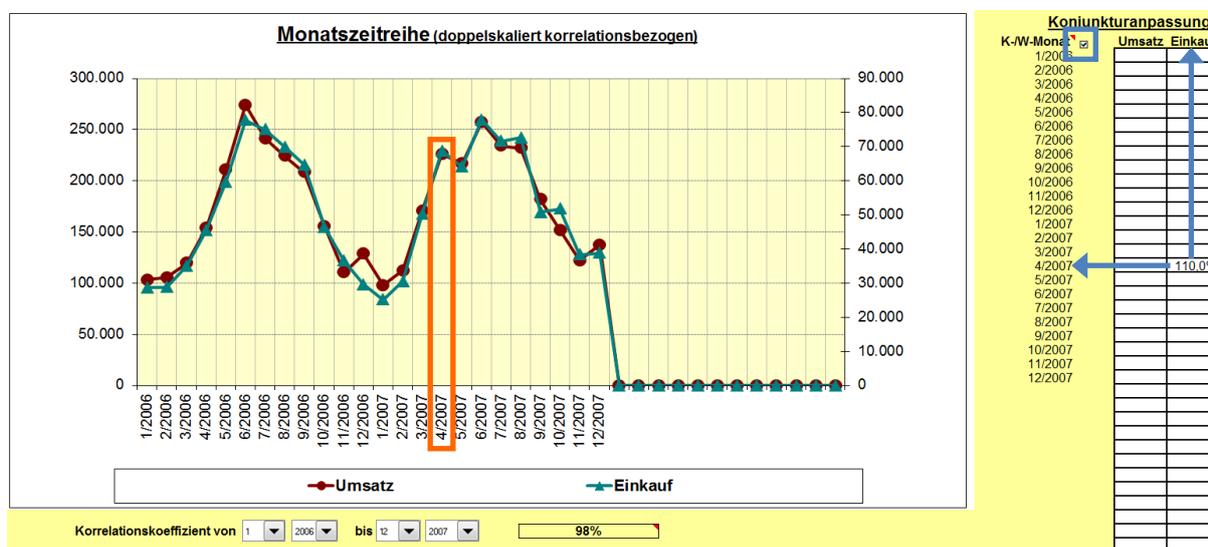
Dieser exemplarischen Gegenüberstellung von monatlichem Warenumsatz und -einkauf eines größeren gastronomischen Betriebs ist der enge Zusammenhang beider Faktoren klar zu entnehmen. Mit 99 % Gleichlauf bestätigt der Korrelationskoeffizient²²⁷ für einen Teilzeitraum unterhalb der Grafik das visualisierte Ergebnis und fällt auch für den gesamten Prüfungszeitraum mit 98 % sehr gut aus. Entgegen Warnungen der Literatur²²⁸ sind Ablösungen infolge von **offenbaren Zuordnungsfehlern** wie im Beispiel der Monate „9-10/2007“ gut zu

²²⁷ Als Maß für den Zusammenhang zweier Datenreihen nimmt der Korrelationskoeffizient Größen von „-1“ für die 100 %-ige Gegenläufigkeit der Grafen über „0“ bei nicht vorhandener Abhängigkeit bis zu „+1“ für eine vollständig positive Gleichrichtung ein.

²²⁸ Wigen, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, 171: „Von großer Bedeutung für die Qualität und Beweiskraft des ZRV ist, dass der Schritt zum zutreffenden ‚Periodenwareneinsatz‘ gelingt.“

erkennen, weil die Interpretation von Zeitreihenanalysen gerade nicht auf einzelne Punkte sondern die (gemeinsame) Entwicklung zweier Grafen abzielt. Ebenso können dem Diagramm prüfungsrelevante **Regelstörungen der Monate „12/2006“ und „4/2007“** entnommen werden.

Nach dem groben Überblick der Quartalszeitreihen zum Verstehen und Differenzieren von systemischen²²⁹ Unternehmensbesonderheiten bzw. -fehlern (s. [Pkt. 4.2.4.](#)) nehmen die Monatszeitreihen eine Zwischenposition für die erforderliche Tiefe (Gründlichkeit) einer Prüfung ein. Vor einer Beurteilung der Schlüssigkeit i.S. des § 158 AO werden entdeckte Einzelauffälligkeiten bis zu möglichen Verursachern zurückverfolgt. Dazu verfügt die SRP-Vorlage BWA über wichtige Funktionen (s.a. [Pkt. 4.2.9.4.](#)) wie die **Konjunkturbereinigung**:



So ergibt die Untersuchung der **Regelstörung „4/2007“** des Beispielsfalls, dass wichtige Waren in dem Monat ausnahmsweise so viel teurer eingekauft werden mussten, dass dadurch ein Gesamtpreisanstieg von 10 % für den Monat resultierte. Über einen entsprechenden **Preisindex** lässt sich dieser unternehmerisch schwerwiegende Sachverhalt umgehend auf Normalverhältnisse berichtigen. Dabei kann die Konjunkturbereinigung zur besseren Nachvollziehbarkeit – besonders von umfangreicheren Korrekturen (s. *das Beispiel „Kostenzeitreihe zum Kraftstoffverbrauch eines Taxiunternehmens“* unter [Pkt. 4.2.9.4.](#)) – über das **Anhakfeld** komplett an- und abgeschaltet werden. Die interaktive Einarbeitung der neuen Erkenntnis zeigt im Beispiel umgehend, dass der außergewöhnliche Umstand die (relative) Ablösung des Warennachkaufs vom Absatz bewirkt hat. Durch das Einpflegen in die Gesamtbildprüfung kann der Störeinfluss problemlos **egalisiert** werden.

²²⁹ Systemische Fehler bzw. Besonderheiten der Datenerfassung und -verarbeitung wie bspw. die verzögerte Aufzeichnung der Einnahmen-Überschuss-Rechnung (§ 4 Abs. 3 EStG) oder falsche Schlüsselungen (z.B. USt-Satz) lassen sich häufig nur oder besser auf stark verdichteten Ebenen wie Quartalen erkennen.

Bei **offensichtlichen Regelablösungen** in einem doppelskaliert größenangepassten Monatszeitreihenvergleich ergeben sich wegen der folgenden Gründe (leicht abgeschwächt) ähnliche Zweifel an der sachlichen Richtigkeit wie bei der Quartalsvariante (*ausführlicher im Abschnitt „offensichtliche Regelablösungen“ unter [Pkt. 4.2.4.](#)*):

- ! die hohe Stabilität des verwendeten Zeitreihenmodells für die Prüfung von Datenreihen auf Abhängigkeit (s. [Pkt. 4.2.](#)),
- ! die bessere Robustheit von Monats- gegenüber Wochensummen²³⁰ als Größenbezug möglicher Zuordnungsfehler in Zeitreihenanalysen,
- ! die gründliche Aufdeckung, Nachverfolgung und mögliche Bereinigung von gesamtbildrelevanten Einzelstörfaktoren wie bspw. von Bestandseinkäufen beim Warenfluss gem. SRP-Zeitreihensystematik (s. [Pkt. 4.2.2.](#)) sowie
- ! die ständigen Prüfungserfahrungen, dass gravierende Regelstörungen in Zeitreihenvergleichen nicht losgelöst von Formal- und anderen materiellen Mängeln vorkommen²³¹ zuzüglich der vielfältigen Optionen, konkrete Erklärungen zügig interaktiv (s. [Pkt. 4.2.9.](#)) daraufhin abzuklären, ob sie ursächlich oder reine Schutzbehauptungen sind.

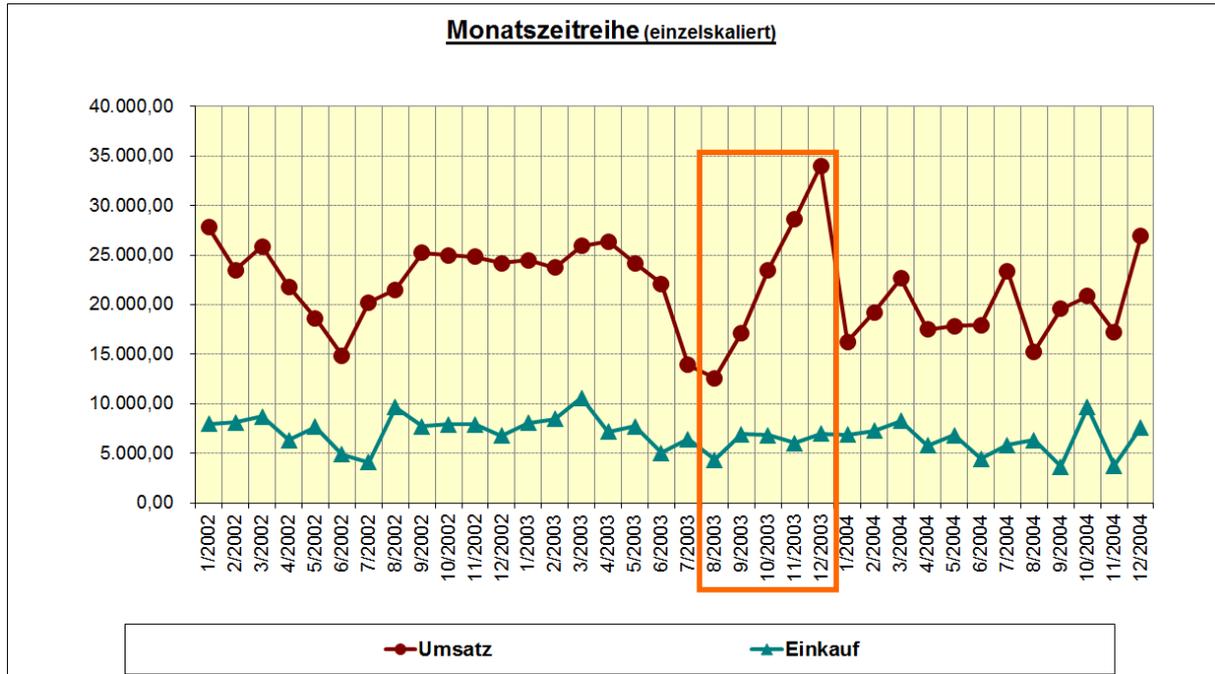
4.2.5.2. Einzelskalierter Zeitreihenvergleich

Mit der zweiten Zeitreihenanalyse auf dem Tabellenblatt „Monatszeitreihen“ werden die beiden Datenreihen in einem einzelskalierten Zeitreihenvergleich gegenübergestellt. Diese Darstellungsform ist dann zielführend, wenn **Übereinstimmung** wie bei der Überprüfung von Datenschnittstellen (s. [Pkt. 5.3.](#)) oder der Kontrolle von Besteuerungsdaten mit Hilfe externer Informationen z.B. vom Zoll erwartet wird.

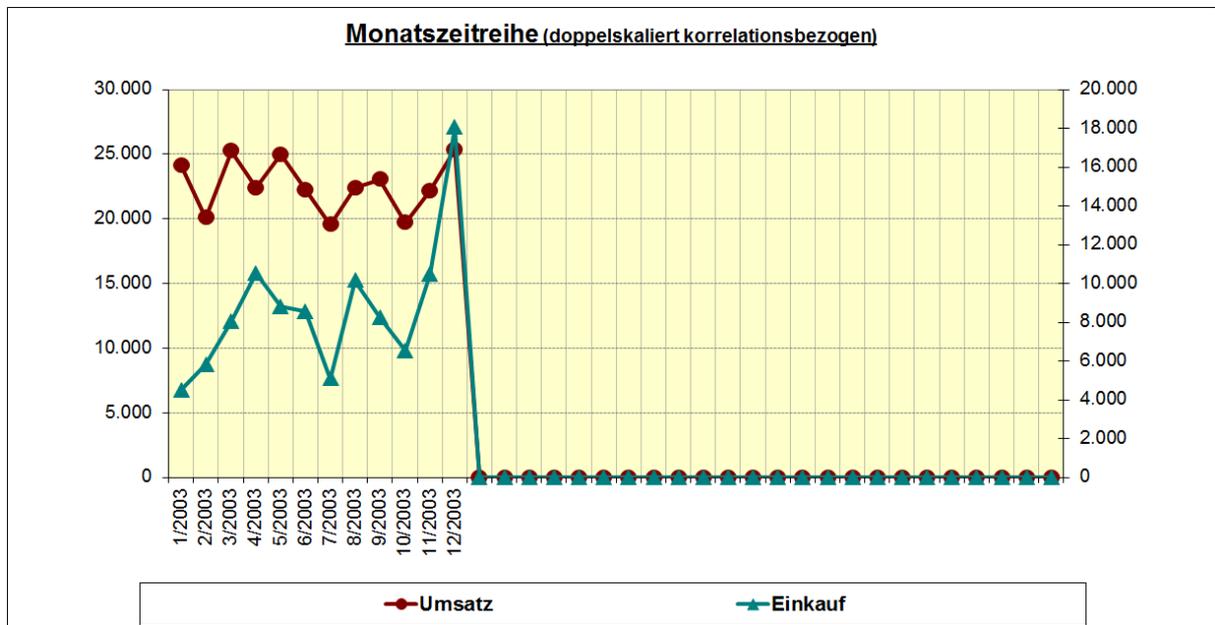
In seltenen Fällen lässt sich bei Regelstörungen in einem Zeitreihenvergleich das Maß an Unschlüssigkeit unter gemeinsamer Skalierung beider Unternehmensfaktoren – trotz der schlechteren Vergleichbarkeit der Einzelentwicklungen – deutlicher herauslesen wie in dem folgenden Beispiel zum Warenfluss eines Restaurants. Besonders der Zeitraum „8-12/2003“ verdeutlicht gem. vorgelegter Besteuerungsgrundlagen eine extreme Absatzzunahme, wofür jedoch keine nennenswerte Erhöhung des Warennachkaufs nötig gewesen sein soll:

²³⁰ Wolenski, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, Stbg 6/2016, 268, mit einem umfangreichen Rechenbeispiel auf S. 269 zur Bedeutung der Größe der jeweiligen Bezugssummen (Wochen, Monate, 10-Wochen-Intervalle, Quartale etc.) auf die Störproblematik.

²³¹ Madauß, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25: „In der Praxis wird sich die Frage, ob ein Zeitreihenvergleich allein die Beweisvermutung des § 158 AO widerlegt, [...] dadurch relativieren, dass die [...] aufgedeckten Unregelmäßigkeiten regelmäßig mit Unregelmäßigkeiten in der Kassenbuchführung einhergehen.“



Da der Geschäftsablauf aufgrund der begrenzten Haltbarkeit von Lebensmitteln unmöglich so stattgefunden haben kann, ist es „unwahrscheinlich [...], dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt“²³².

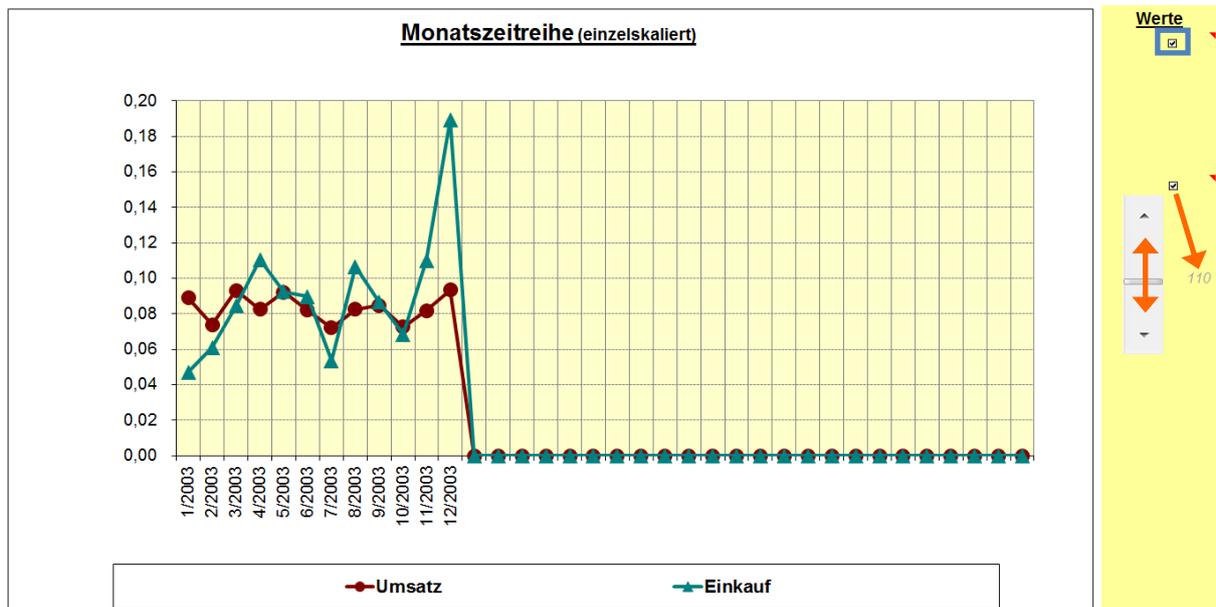


Mitunter verzerrt ein unzusammenhängender²³³ Ausreißer einer Datenreihe die sekundäre Y-Achse wie in folgender Abb. der Wareneinkauf „12/2003“ eines Dönerimbisses. Hierdurch wird die Vergleichbarkeit der Entwicklungen beider Datenreihen trotz größenbezogener Y-Achsen verhindert. Neben der Nachverfolgung und Berichtigung der ursächlichen Sachver-

²³² Anwendungserlass zu § 158 AO.

²³³ Sofern beide Datenreihen starke Schwankungen in ihren Monatssummen durchlaufen, dabei aber die erwartete schlüssige Reaktion aufeinander zeigen, wird die Größenanpassung via Doppelskalierung nicht gestört.

halte gem. SRP-Prüfungssystematik (s. [Pkt. 4.2.2.](#)) bietet die SRP-Vorlage BWA als Schnelllösung eine variabel-relative Form des einzelskalierten Zeitreihenvergleichs an:



Über die **Umschaltfunktion** werden die Monatswerte für Warenumsatz und -einkauf als Anteile der jeweiligen Jahressummen ausgegeben. Nun kann die erste Datenreihe für den Wareneinkauf – **nach Setzen des Hakens** – **mittels Schieberegler** auf das Größenniveau des zweiten Grafen gezogen werden. Im Beispielsfall offenbart sich auf diese Weise ohne langwierige Vertiefungsmaßnahmen, dass die Regelstörungen (nahezu) das komplette Jahr betreffen. Ein zeitlicher Versatz der gesamten Datenreihe „Wareneinkauf“ infolge der verzögerten Erfassung i.R. der Einnahmen-Überschuss-Rechnung nach § 4 Abs. 3 EStG lässt sich weder in der Monats- noch in der – gem. SRP-Zeitreihe-systematik vorgeschalteten, übersichtlicheren – Quartalsansicht erkennen. In der Konsequenz kann die Unschlüssigkeit i.S. des § 158 AO nur auf materielle Unrichtigkeit der Besteuerungsgrundlagen oder so gravierende Ordnungsmängel zurückgehen, dass diese für sich bereits die Beweiskraft der Aufzeichnungen widerlegen.

4.2.5.3. Verhältniszeitreihe (z.B. Aufschlagsatzzeitreihen)

Die Zeitreihenanalyse zur Entwicklung der monatlichen Verhältnisgröße²³⁴ wird unter dem Thema „Schätzung im Besteuerungsverfahren (Quantilsschätzung)“ (s. [Pkt. 4.8.](#)) beschrieben, weil sie zur Steigerung der Transparenz einer Quantilsschätzung und zur sachgerechten Anpassung an den Einzelfall beitragen soll.

I.R. einer Plausibilitätsprüfung mit SRP, wie sie häufig am Gesamtbild des Warenflusses ansetzt, hat die Entwicklung der Aufschlagsätze keine vorrangige Bedeutung (zur Bedeutung

²³⁴ Der Verhältnisfaktor wird i.S. eines Aufschlagsatzes ermittelt mit: $\frac{\text{Datenreihe 2} - \text{Datenreihe 1}}{\text{Datenreihe 1}}$ und prozentual ausgegeben.

des Richtsatzvergleichs auf dem Tabellenblatt „Jahressummen“ s. ebenso unter [Pkt. 4.2.3.](#)). Zum einen wird der externe Vergleich wegen seiner begrenzten Übertragbarkeit auf den Einzelfall in der AO-Kommentierung als „größte Methode“²³⁵ bezeichnet. Zum anderen zeigen Aufschlagsatzzeitreihen²³⁶ der Prüfungspraxis im Gegensatz zu geeigneteren Modellen eine deutliche Anfälligkeit gegenüber eher alltäglichen Störeinflüssen (s. *den Modellvergleich im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)*).

4.2.6. Gleitschlittenzeitreihen²³⁷

Gleitende Zeitreihen können die Robustheit längerer Vergleichsperioden mit der Detailliertheit von kürzeren verbinden, indem sie z.B. ein für Störeinflüsse wenig anfälliges 2-Monats-Intervall immer nur um einen Tag forttragen: Auf die Startperiode „1.1. bis 28.2.“ folgt dann „2.1. bis 1.3.“, „3.1. bis 2.3.“, 4.1. bis 3.3.“ etc. Einerseits werden dadurch nur Regelstörungen angezeigt, die so schwerwiegend sind, dass sie einen 2-Monatswert verzerren können.²³⁸ Andererseits bewirken sowohl der Einstieg als auch der Ausstieg eines ursächlichen Sondervorgangs jeweils eine abrupte Veränderung des betreffenden Grafen, indem bspw. ein hochwertiger Bestandseinkauf am 13.4. den 2-Monatsaufschlagsatz erstmalig für die Periode „14.2. bis **13.4.**“ (plötzlich) absenkt. In den Folgeintervallen „15.2. bis 14.4.“, „16.2. bis 15.4.“ usw. beinhaltet die 2-Monatsgröße weiterhin den Sondersachverhalt, so dass der Aufschlagsatz plateauartig niedrig bleibt, bis er sich mit dem Betrachtungszeitraum „14.4. bis 13.6.“ schlagartig erholt, denn der Bestandseinkauf wird nicht mehr mitgezählt.

Aufgrund dieser besonderen Eigenschaft der Gleitschlittenzeitreihen werden sie i.R. der SRP zur weiteren **Eingrenzung von Regelstörungen** eingesetzt, welche mit doppelskaliert größenangepassten Zeitreihenvergleichen entdeckt worden sind (s. *die SRP-Zeitreihensystematik unter [Pkt. 4.2.2.](#)*). Als [Standardeinstellung des Gleitschlittens](#) hat sich aus den ständigen Prüfungserfahrungen das Intervall „2 Monate“ herauskristallisiert. Über das

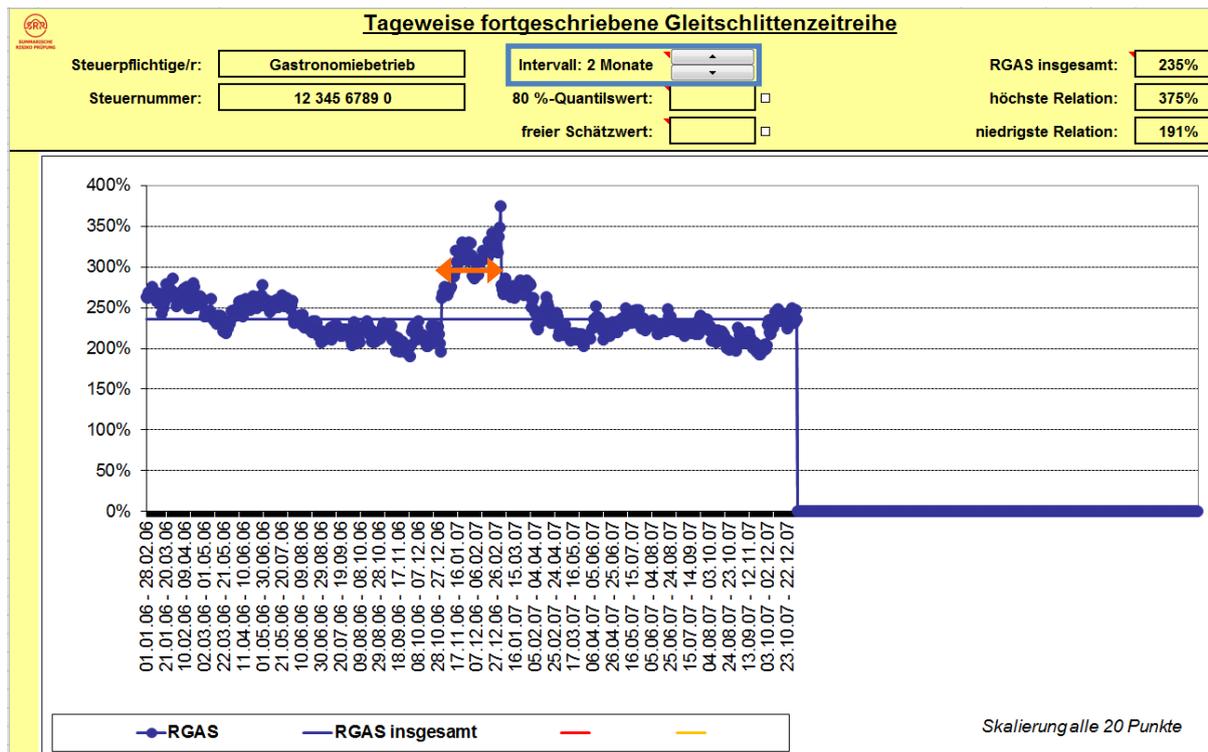
²³⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 17: „Als größte Methode zur Beanstandung der sachlichen Richtigkeit des Buchführungsergebnisses erweist sich der äußere Betriebsvergleich [...] Vergleichen mit anderen Betrieben kommt regelmäßig ein geringerer Vergleichswert als inneren Betriebsvergleichen [...] zu“.

²³⁶ Auch dem BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743) liegt mit der „Gleitschlittenzeitreihe“ eine Form der Aufschlagsatzzeitreihe zu Grunde, weshalb die erläuternde Literatur der betreffenden RiBFH z.T. die Anwendbarkeit auf entsprechende Zeitreihenvarianten beschränkt [Nöcker (RiBFH), „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548].

²³⁷ Harle/Olles, „Die moderne Betriebsprüfung“, NWB Verlag, 3. Auflage 2017, Rz. 1366 ff.

²³⁸ Wolenski, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, Stbg 6/2016, 268, zeigt auf S. 269 mit einem umfangreichen Rechenbeispiel zur 10-Wochen-Gleitschlittenzeitreihe des BFH-Zeitreihenurteils (v. 25.3.2015, BStBl II. 2015, 743), wie entscheidend die Größe der Bezugssummen der gewählten Vergleichsperiode (Wochen, Monate, 10-Wochen-Perioden, Quartale etc.) für den Einfluss von Störfaktoren ist.

Drehfeld sind jedoch auch die Alternativen „3 Monate“, „10 Wochen“, „6 Wochen“, „1 Monat“ und „14 Tage“ auswählbar:



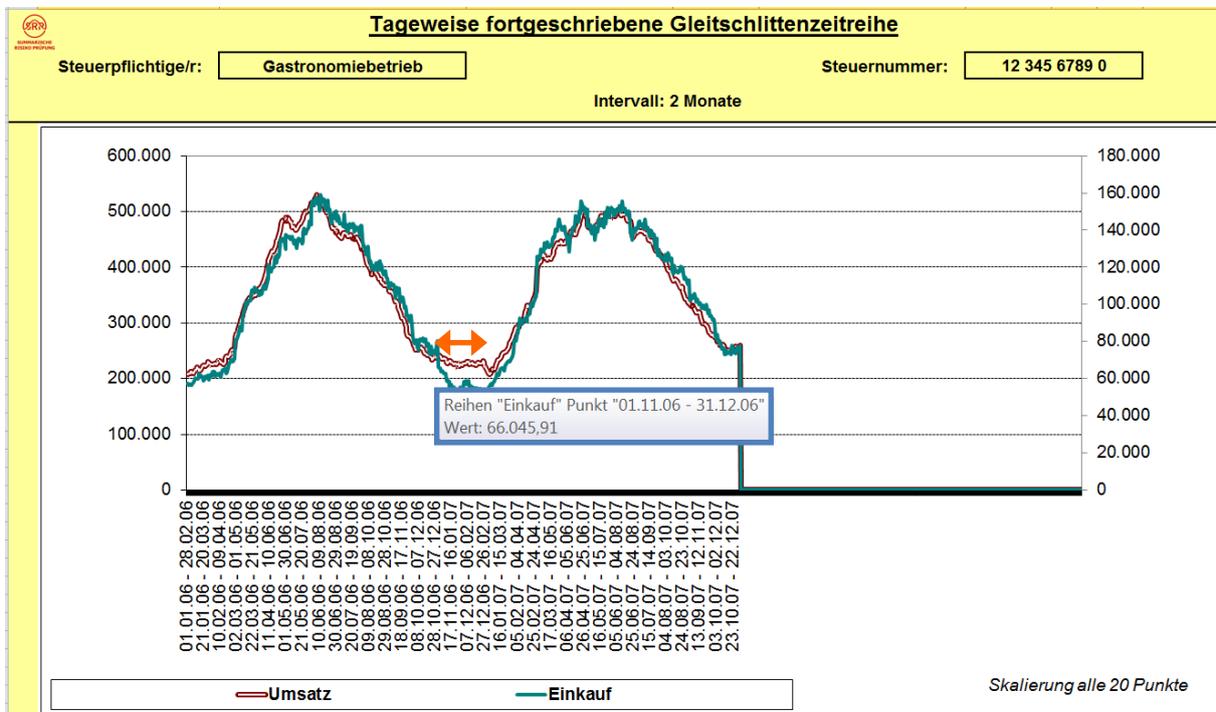
Im Beispiel sind die **zwei abrupten Veränderungen des Aufschlagsatzes**, die durch einen einzelnen Tag – und damit vermutlich einen Einzelsachverhalt – verursacht werden, gut zu erkennen. Wenn die Maus im Diagramm entweder auf den Plateaustart oder das -ende gelegt wird („Mouse-over“), erscheint ein Beschreibungsfähnchen für den jeweiligen Punkt. Damit lässt sich die Verzerrung hier auf das Erstintervall „1.11. bis 31.12.2006“ konkretisieren, so dass der ursächliche Vorgang am 31.12.2006 zu suchen ist.

Über Filter- und Suchfunktionen können nun die ursächlichen „Regelstörer“ zügig identifiziert und geprüft werden. In diesem Fall lässt sich auf dem Weg in den verbleibenden Daten eines Tages unproblematisch ein außerordentlich hoher Negativbetrag erkennen. Nach Belegprüfung stellt sich der Vorgang als Jahresbonus heraus, der von den regelmäßigen Wareneinkäufen nicht über ein Extrakonto abgegrenzt worden ist. Im Ergebnis ist zu prüfen, ob im Folgejahr eine entsprechende Erfolgsvergütung gezahlt worden ist, denn die Gleitschlittenzreihe zeigt keine vergleichbare Entwicklung. Außerdem muss der Vorgang für die Beurteilung der Plausibilität des Warenflusses i.S. des § 158 AO interaktiv (s. [Pkt. 4.2.9.1.](#)) korrigiert werden, weil die zeitliche Zuordnung zum 31.12. operativ falsch und der gesamtbildrelevante Störeinfluss offensichtlich ist.

Der **doppelskaliert größenangepasste Zeitreihenvergleich per Gleitschlitten** zu beiden Datenreihen grenzt aufgefundene Regelstörungen auf einzelne Tage und die jeweilige Datenreihe ein. Damit führt er i.R. der SRP-Zeitreihensystematik (s. [Pkt. 4.2.2.](#)) noch

schneller zu prüfungsrelevanten Sachverhalten, erfordert allerdings aufgrund der Komplexität der Darstellung mehr Analyseerfahrung.

Zu dem Beispielsfall von oben sieht der differenzierende Zeitreihenvergleich in gleitender 2-Monatsform wie folgt aus:



Im Zeitfenster um den Monat „12/2006“, in welchem gem. doppelskaliertem Monatszeitreihenvergleich eine Regelstörung aufgefallen ist, fällt lt. **Mouse-over-Fähnchen** die Datenreihe „Einkauf“ im Intervall „1.11. bis 31.12.“ abrupt ab. Die korrespondierende Gegenbewegung beim Ausschleichen des Tages mit dem Sondersachverhalt ist hinter dem Anzeigefähnchen verborgen.

Die Konsequenzen aus der noch etwas konkreteren Identifikation des gesamtbildstörenden Einzelvorgangs erfolgen analog der Vorgehensweise bei der Gleitschlittenzitreihe zum Aufschlagsatz (s. den Abschnitt „Eingrenzung von Regelstörungen“ unter diesem Pkt.).

Für die Plausibilitätsprüfung i.S. des § 158 AO mit SRP werden die Gleitschlittenzitreihen nicht herangezogen, weil die Aufschlagsatzvariante die übliche Anfälligkeit gegenüber alltäglichen Einflussfaktoren zeigt (s. den Vergleich von drei Zeitreihentypen im Abschnitt „Methodenkenntnis“ unter Pkt. 4.2.) und die doppelskalierte Form von gleitenden Zeitreihenvergleichen schwer zu lesen ist. Aus diesem Grund wirkt die interaktive Konjunkturbereinigung des Tabellenblatts „Monatszeitreihen“ (s. Pkt. 4.2.5.1.) nicht auf die Gleitschlittenmodelle durch.

Die Anhak-Optionen, mit welchen sich die Schätzungsparameter „80 %-Quantil“ und „freier Schätzwert“ in der Aufschlagsatz-Gleitschlittenzitreihe einblenden lassen (s. die obere Abb. zu diesem Pkt.), dienen ausschließlich der optischen Überprüfung, ob eine mit SRP

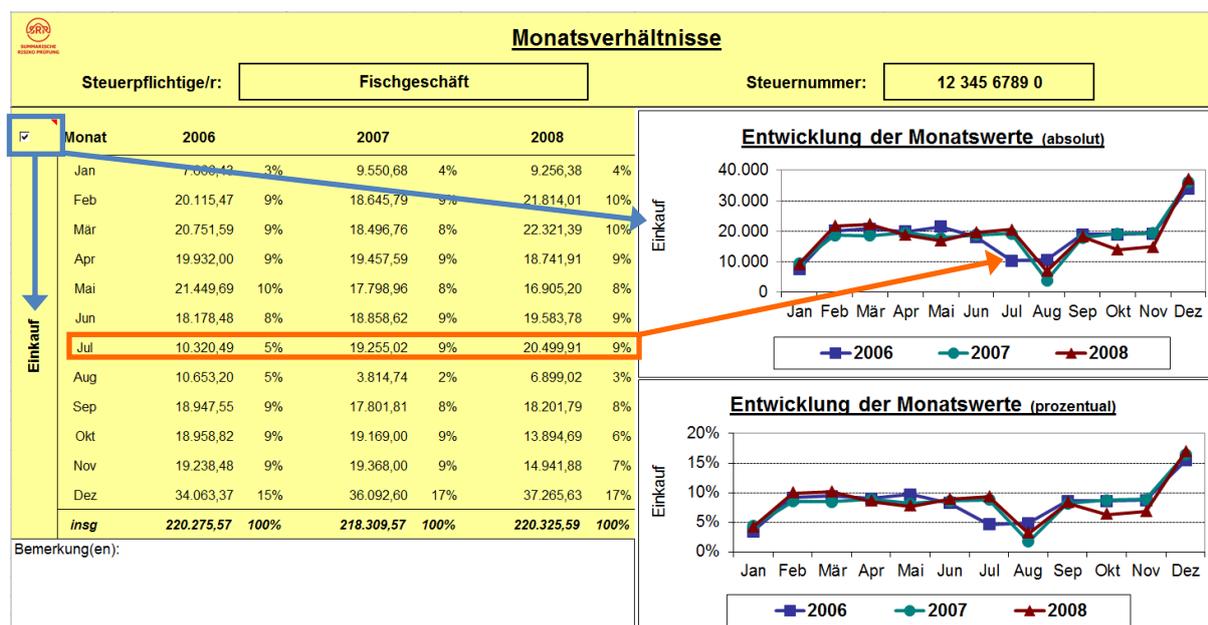
festgelegte Schätzungshöhe am Maßstab der innerbetrieblichen Entwicklung sachgerecht erscheint und wird *unter Pkt. 4.8.* beschrieben.

4.2.7. Entwicklungsvergleich der einzelnen Datenreihen

Eine **gesonderte Entwicklungsbetrachtung der einzelnen Datenreihen** durch eine Gegenüberstellung über mehrere Jahre kann aus Besteuerungsdaten weitere Informationen zur zügigen Identifikation prüfungsrelevanter Sachverhalte oder direkte Erkenntnisse für eine Plausibilitätsbeurteilung beisteuern. Dazu werden auf den Tabellenblättern „Monatsverhältnisse“ und „Wochentagswerte“ der SRP-Vorlage BWA die Jahresgrafiken zur betreffenden Datenreihe als Monatssummen bzw. als Durchschnitte für die Wochentage mit ihren Werten und prozentual gegenübergestellt.

Für eine Vertiefung der Ergebnisse bspw. zum Differenzieren von generellen Trends wie einer jahresweisen Steigerung der Wochentags-Mittelwerte gegenüber der Verzerrung durch einzelne Extremwerte bieten sich die Häufigkeitsanalysen der Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ in der SRP-Vorlage Stochastik an (s. [Pkt. 4.3.3.](#)).

Die – in der Unternehmensplanung und -beratung geschätzte – aussagekräftige Stabilität von betrieblichen Monatssummen (s. [Pkt. 4.2.5.](#)) kann für unterschiedliche Formen von Zeitreihenanalysen genutzt werden. Im Unterschied zur Prüfung, ob zwei Betriebsfaktoren eine handlungslogisch erwartete Abhängigkeit zeigen, stellt die **jahresweise Gegenüberstellung der monatlichen Entwicklung einer Betriebsgröße** außerordentliche Veränderungen gegenüber Saisonschwankungen häufig deutlicher heraus. Über ein [Anhakfeld](#) lässt sich die gewünschte Datenreihe auswählen und wird in der Tabelle sowie in den Diagrammen angezeigt:



Diese Grafik zu einem Fischgeschäft lässt den üblichen Wareneinkaufsrhythmus über das Jahr gut erkennen, indem die drei Grafen weitgehend deckungsgleich verlaufen. Vorhandene Ablösungen wie z.B. für das erste Jahr im Mai oder für das letzte in der Phase „Oktober bis November“ können auf konjunkturelle Einflüsse zurückgehen und sollten sich entsprechend in abhängigen Betriebsfaktoren – also hier dem Umsatz – wiederfinden.

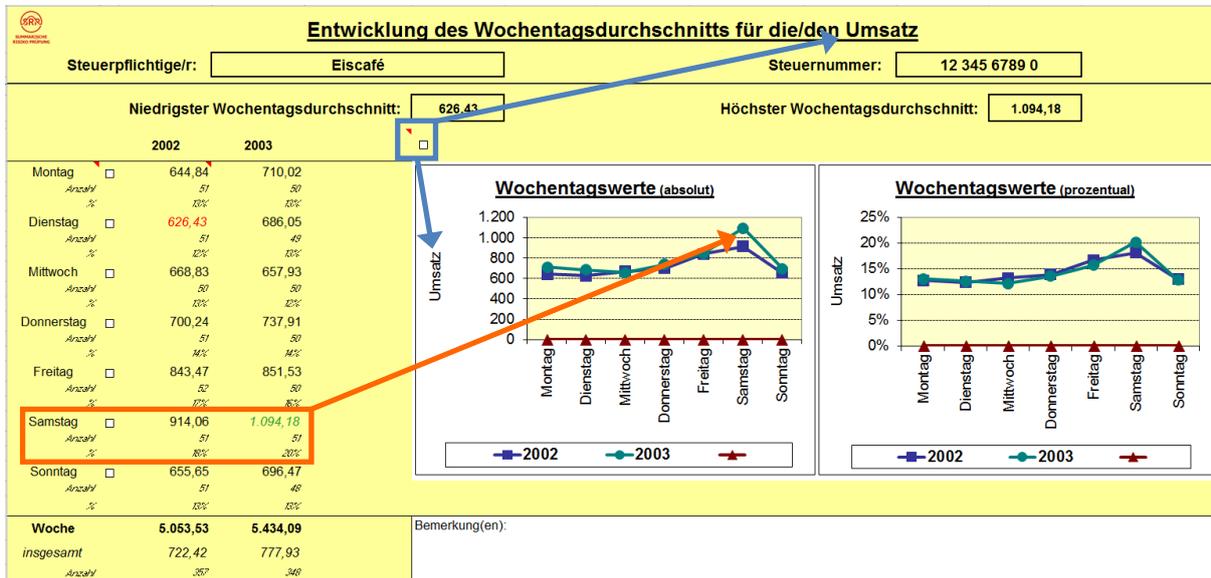
Auffallend ist jedoch die **andersartige Absenkung in den Sommermonaten „Juli bis August“**. Als mögliche Ursachen einer Regelstörung kommen einzelne verzerrende Sachverhalte sowie kurzfristige Strukturveränderungen²³⁹ im Unternehmen in Betracht. Mit Hilfe der SRP-Zeitreihensystematik lassen sich außergewöhnliche Sondervorgänge als Störeinflüsse identifizieren oder ausschließen (s. [Pkt. 4.2.2.](#)). Da zu diesem Beispiel keine Einzelverursacher auffindbar sind, muss der Auslöser in einer grundsätzlichen Ablaufbesonderheit mit einer Abweichung im ersten Jahr liegen. Aus der entsprechenden Nachfrage im Unternehmen geht hervor, dass die Betriebsferien üblicherweise komplett im August stattfinden und in 2006 ausnahmsweise zur Hälfte in den Juli vorverlegt wurden. Diese Erkenntnis zu den betriebspezifischen Verhältnissen trägt in der Folge zu einer qualifizierten Plausibilitätsbeurteilung i.S. des § 158 AO bei.

Im Hinblick auf die dokumentierte Absatzsituation bei bargeldintensiven Betrieben sowie bei der Untersuchung einer Vielzahl von Kostenpositionen auf prüfungsrelevante Sonderfälle können aus dem **Jahresvergleich der durchschnittlichen Wochentagswerte** wichtige Informationen hervorgehen. Weil arithmetische Mittelwerte sehr anfällig gegenüber Verzerrungen durch einzelne Ausreißer sind²⁴⁰, wird die Anzahl der vorhandenen Wochentage mit den Daten zum Durchschnitt ausgegeben. Auch bei dieser Auswertung kann die zu untersuchende Datenreihe über ein **Anhakfeld** ausgewählt werden und erscheint danach in der Überschrift und den Diagrammen.

²³⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, 1. Leitsatz: „Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist. Es darf zudem im maßgebenden Zeitraum nicht zu solchen Änderungen in der Betriebsstruktur gekommen sein, die – nicht anderweitig behebbar – wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringen.“ Geeignete Zeitreihentypen wie hier im Beispiel oder der doppelskaliert größenangepasste Zeitreihenvergleich für Abhängigkeitsuntersuchungen lassen solche Strukturänderungen jedoch als sinnvolle Vertiefungsfelder eindeutig erkennen (s.a. den Abschnitt „Methodenkennntnis“ unter Pkt. 4.2.).

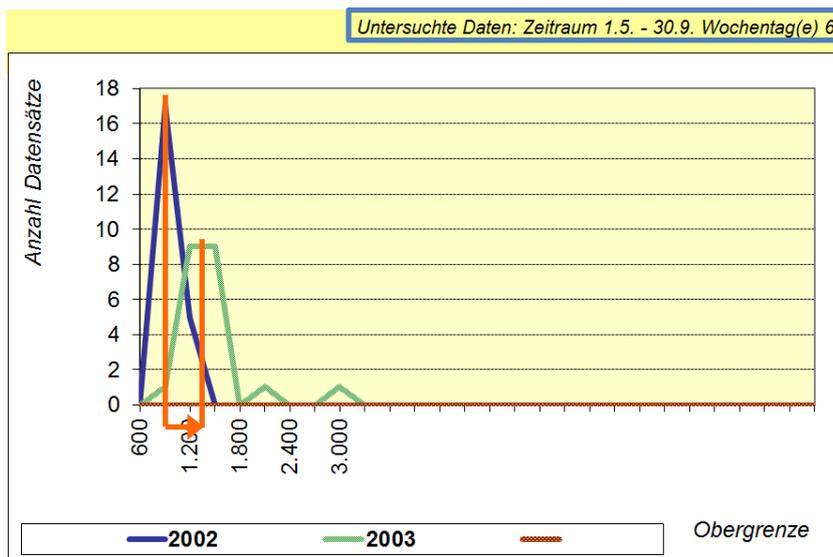
²⁴⁰ Arithmetische Kennzahlen wie Summen, Mittelwerte etc. sind infolge der Rechenoperation anfällig für einzelne Ausreißerwerte: Wenn z.B. die Werte „10“, „10“, „11“, „12“, „12“, „13“, „15“, „15“, „18“ und „65“ durch ihren Mittelwert repräsentiert werden sollen, ergibt sich „18,1“. Rangkennzahlen, welche die Daten „ranken“ und dann einen absoluten oder relativen (Prozentränge = Quantile oder Bruchränge = Fraktile) Rang auswählen, bilden bei asymmetrischen Verteilungen, wie sie häufig in Wirtschaftsdaten vorkommen, meistens bessere Stellvertreter. So trifft der Median (50 %-Rang) in diesem Fall mit dem Wert „12,5“ [zwischen Rang „5“ und „6“ wird interpoliert (zwischenberechnet)] den Großteil der Daten deutlich besser.

Für ein beispielhaftes Eiscafé ist die sehr ähnliche unterwöchige Nachfragesituation über die beiden Jahre in Form der Kongruenz (Deckungsgleichheit) beider Grafen ablesbar. Dagegen sticht die einseitige **Verbesserung der Umsatzlage für die Samstag des zweiten Jahres** hervor:



Der außerordentliche Anstieg der Samstagserlöse um fast 20 % mit ca. € 180 bedeutet eine klare Störung der erwarteten Regel „gleichmäßige Nachfrage“. Als mögliche Gründe für diese Auffälligkeit kommen einzelne Verzerrer als besonderer Absatz oder Erfassungsfehler sowie eine generelle Erhöhung in Frage.

Trotz der Trennung von regel- und musterbasierten Analysen in den beiden SRP-Vorlagen, wodurch die Bedienung erleichtert werden soll, können bei der vertiefenden Analyse derartiger Auffälligkeiten vor allem Häufigkeitsauszählungen helfen. In diesem Fall offenbart der Vergleich beider Jahresverteilungen zu den **Samstagserlösen der Saison** auf dem Tabellenblatt „Schichtungsver-



teilung“ der SRP-Vorlage Stochastik (s. [Pkt. 4.3.3.](#)) unmissverständlich, dass die gesamte Kurve der Samstag aus 2002 im Folgejahr auf der X-Achse nach rechts zu den höheren Werten gewandert ist. Der grünen Häufigkeitsverteilung für das Jahr „2003“ sind zwar mit den beiden Extragipfeln rechts auch zwei besonders ertragreiche Tage zu entnehmen. Der

Niveaustieg der Häufungszone von ca. € 900 auf ca. € 1.250 ist jedoch hauptsächlich für die in den Wochentagsdurchschnitten erkannte Nachfrageverbesserung.

Aus Plausibilitätssicht ist das Gesamtbild²⁴¹ danach zu beurteilen, ob es glaubhaft ist, dass die Kundennachfrage von einem Jahr auf das nächste ausschließlich an einem Wochentag wesentlich zugenommen hat, oder ob es eher „unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt“²⁴². Mögliche einzelbetriebliche Begründungen für diese außergewöhnliche Beobachtung wie bspw. regelmäßige Sonderverkaufsaktionen an Samstagen ab 2003 können somit i.R. der gemeinsamen Sachaufklärungsverantwortung²⁴³ konkret hinterfragt werden.

4.2.8. Rückrechnung der monatlichen Warenbestände

Die rechnerische Rekonstruktion – nicht aufgezeichneter – unterjähriger Warenbestände stellt in der SRP keine eigenständige Verprobung dar, sondern soll der Verdeutlichung von Regelstörungen dienen, die mit Hilfe von doppelskalierten Zeitreihenvergleichen aufgedeckt wurden (zur unterschiedlichen Eignung bestimmter Zeitreihenvarianten für die Abhängigkeits- und damit Plausibilitätsprüfung s. den Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) und trotz SRP-Zeitreihensystematik (s. [Pkt. 4.2.2.](#)) und interaktiver Optionen zum Einpflegen von konkreten Erklärungsversuchen (s. [Pkt. 4.2.9.](#)) nicht aufgeklärt werden konnten.

Um die Auswirkungen eines un schlüssigen Wareneinkaufverhaltens greifbarer zu veranschaulichen, können die **unterjährigen Warenbestände anhand der Periodendaten rückgerechnet** werden, sofern sie nicht aufgezeichnet worden sind. In der SRP-Vorlage BWA werden dazu die Datenreihen „Warenumsatz“ und „Wareneinkauf“ inkl. interaktiver (s. [Pkt. 4.2.9.1.](#)) Zuordnungskorrekturen sowie die Eingaben zu Warenanfangsbestand (*Tabellenblatt „Bestandsentwicklung“*), jährlichen Bestandsveränderungen und Entnahmen (*beides auf dem Tabellenblatt „Jahreszahlen“*) verwendet. Unter Annahme eines konstanten Aufschlagsatzes²⁴⁴ wird für jeden Monat der erforderliche Wareneinsatz²⁴⁵ berechnet und

²⁴¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

²⁴² Anwendungserlass zu § 158 AO.

²⁴³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

²⁴⁴ Damit teilt die rechnerische Rekonstruktion unterjähriger Warenbestände die größere Empfindlichkeit von Aufschlagsatzzeitreihen gegenüber Störeinflüssen im Vergleich zu den doppelskaliert größenangepassten Zeitreihenvergleichen (s. den direkten Vergleich im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)). Insofern

dem Periodennachkauf gegenübergestellt. Mit der Differenz wird der Warenanfangsbestand monatlich fortgeführt.

Monat	Umsatz	Entnahme	notwendiger Einsatz	Einkauf	Bestandsveränderung	laufender Bestand
	2006	<input checked="" type="checkbox"/>	RGAS: 244,27%		Anfangsbestand:	22.620,00
Jan 2006	103.316,36	795,61	30.806,27	28.669,66	-2.136,61	20.483,39
Feb 2006	105.679,39	795,61	31.492,67	28.849,07	-2.643,60	17.839,79
Mär 2006	119.460,88	795,61	35.495,82	35.092,60	-403,22	17.436,57
Apr 2006	154.220,90	795,61	45.592,69	45.561,73	-30,96	17.405,61
Mai 2006	210.792,54	795,61	62.025,25	59.712,01	-2.313,24	15.092,38
Jun 2006	274.183,18	795,61	80.438,55	77.790,74	-2.647,81	12.444,57
Jul 2006	241.253,15	795,61	70.873,25	75.018,82	4.145,57	16.590,15
Aug 2006	224.315,47	795,61	65.953,30	73.903,05	7.949,75	24.539,90
Sep 2006	208.366,35	795,61	61.320,50	64.639,40	3.318,90	27.858,79
Okt 2006	155.559,82	795,61	45.981,61	46.412,66	431,05	28.289,85
Nov 2006	110.351,57	795,61	32.849,81	36.503,51	3.653,70	31.943,54
Dez 2006	128.789,24	795,61	38.205,46	27.627,92	-10.577,54	21.366,00
<i>insgesamt</i>	<i>2.036.288,85</i>	<i>9.547,30</i>	<i>601.035,17</i>	<i>599.781,17</i>	<i>-1.254,00</i>	
	2007		RGAS: 246,53%			

In der exemplarischen Tabelle wurde die **Anhak-Option** „jährlich gesonderte Betrachtung“ ausgewählt, wodurch wegen geänderter Bedingungen für jedes Jahr ein eigener Aufschlagsatz zu Grunde gelegt wird. Bleibt das Feld hingegen deaktiviert, wird für alle Jahre mit einem einheitlichen Gesamtaufschlagsatz²⁴⁶ gerechnet. Bei bilanzierenden Stpfl. werden die Jahresendbestände der Zwischenjahre im zweiten Fall nicht mit den ausgewiesenen übereinstimmen, falls die jährlichen Aufschlagsätze abweichen.

Als **mögliche Implausibilitäten der rückgerechneten Warenbestände** können (rechnerisch) negative oder unrealistisch hohe Zwischenstände resultieren. Dazu kann ein maximal möglicher Lagerbestand rechts oberhalb der ersten Grafik des Tabellenblatts „Bestandsentwicklung“ eingetragen werden, damit das zeitweise (rechnerische) Überschreiten im Diagramm offensichtlich wird.

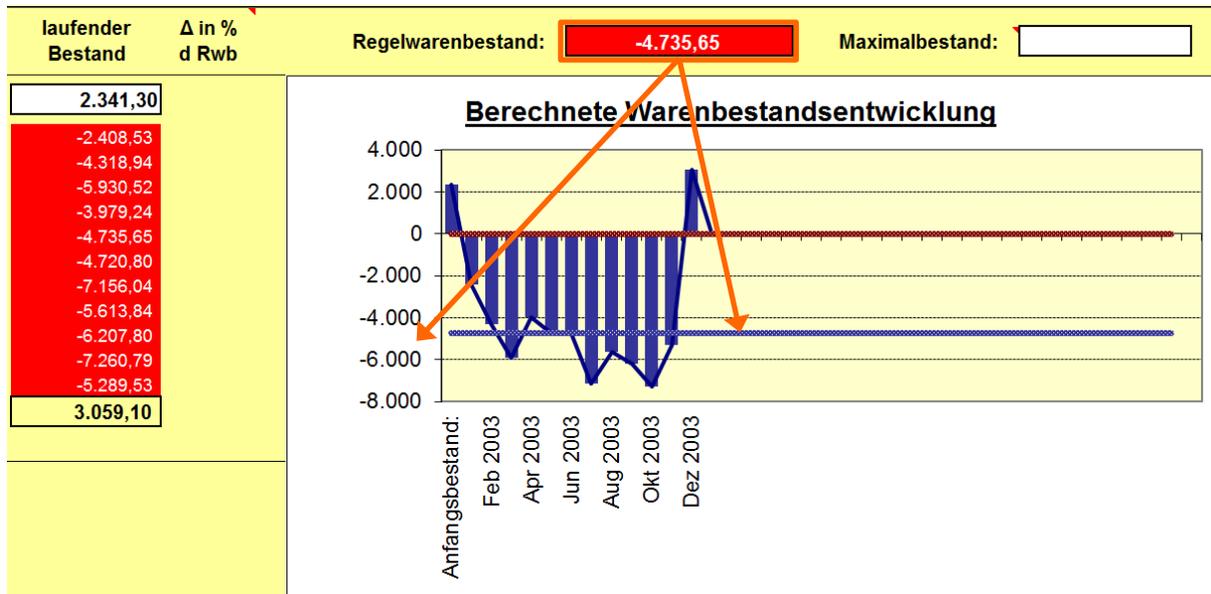
Das nachfolgende Beispiel eines Dönerimbisses offenbart im Prüfungsjahr permanent negative Zwischenbestände des Warenlagers zu den Monatsenden, welche in der Tabelle farblich hervorgehoben werden. Deshalb ergibt sogar das **Lagerrepräsentativ**, welches als

trifft hier die Warnung aus dem BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743, 1. Leitsatz) zu: „Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist. Es darf zudem im maßgebenden Zeitraum nicht zu solchen Änderungen in der Betriebsstruktur gekommen sein, die – nicht anderweitig behebbar – wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringen.“

²⁴⁵ Berechnung des monatlichen Wareneinsatzes nach der Formel: $\frac{\text{Warenumsatz}}{(\text{Aufschlagsatz}+100\%)} + \text{Entnahmen}$.

²⁴⁶ Der Gesamtaufschlagsatz für den auf dem Tabellenblatt „Vergleichsdaten“ eingegebenen Prüfungszeitraum wird auch auf dem Tabellenblatt „Gleitschlitzenzeitreihen“ oben rechts sowie in dem dazugehörigen Aufschlagsatzdiagramm ausgegeben (s.a. die erste Abb. unter Pkt. 4.2.6.).

Median (50 %-Rang) aller eingegebenen zuzüglich rückgerechneten Warenbestände ermittelt wird, einen **Negativwert**:



Diese Perspektive zu den Bestimmungsaufzeichnungen des Warenflusses veranschaulicht gegenständlicher, dass die bereits in den Monatszeitreihen aufgefallenen Einkaufsdaten hinsichtlich ihrer Plausibilität nicht nachvollziehbar sind. So hat dort der unzusammenhängende²⁴⁷ Ausreißer nach oben „Dez 2003“ die Vergleichbarkeit von Warenumsatz und -einkauf trotz größenanpassender Doppelskalierung verhindert (s. [Pkt. 4.2.5.2.](#)). Durch die Umstellung der Monatswerte auf Anteile der jeweiligen Jahressummen im Zeitreihenvergleich mit anschließend geringfügiger Niveaueinstellung der Grafen ließ sich aus dem Diagramm entnehmen, dass der Zusammenhang der Warenflussdaten über das gesamte Jahr nicht gegeben ist²⁴⁸. Aus Sicht der unterjährig rückgerechneten Warenbestände zeigt sich die bekannte Unschlüssigkeit darin, dass über das Jahr stets zu wenig Ware für den erzielten Umsatz vorhanden war, weil sie erst zum Jahresende konzentriert erfasst worden ist.

Solange sich derart unschlüssige Zwischenstände des Warenlagers nicht durch die gemeinsame²⁴⁹ interaktive (s. [Pkt. 4.2.9.](#)) Korrektur der zeitlichen Zuordnung gesamtbildrelevanter Wareneinkäufe (weitgehend) auflösen lassen, geben sie Anlass für berechtigte Zweifel an der sachlichen Richtigkeit von Besteuerungsgrundlagen. Gem. umfangreicher Prü-

²⁴⁷ Sofern beide Datenreihen starke Schwankungen in ihren Monatssummen durchlaufen, dabei aber die erwartete schlüssige Reaktion aufeinander zeigen, wird die Größenanpassung via Doppelskalierung nicht gestört.

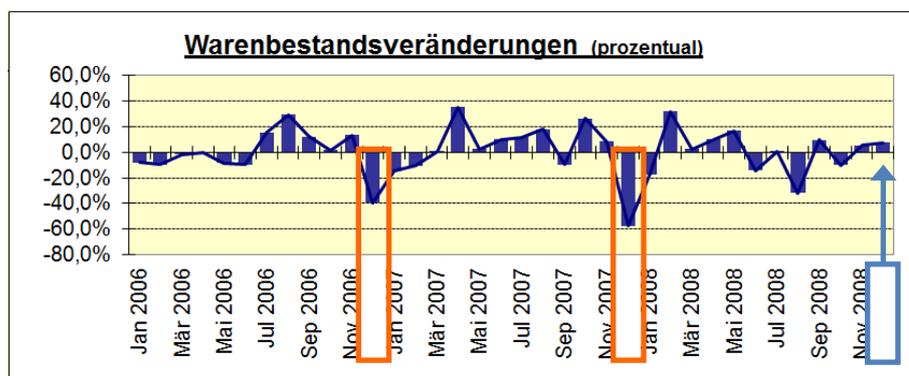
²⁴⁸ Mit Hilfe der interaktiven Optionen (s. [Pkt. 4.2.9.](#)) der SRP-Vorlage BWA konnte zudem ausgeschlossen werden, dass die Aufzeichnungsweise „Einnahmen-Überschuss-Rechnung“ nach § 4 Abs. 3 EStG (alleine) ursächlich für die Regelstörungen ist.

²⁴⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen“.

fungserfahrungen mit SRP gehen unschlüssige rechnerische Warenzwischenstände mit Regelstörungen des deutlich robusteren doppelskalierten Zeitreihenvergleichs (*zur unterschiedlichen Eignung bestimmter Zeitreihenvarianten für die Plausibilitätsprüfung s. den Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)*) einher, so dass die Schlüssigkeitsbeurteilung i.S. des § 158 AO auf der **mehrperspektivisch abgesicherten** Analyse des Warenflusses lt. Besteuerungsaufzeichnungen basiert.

Sofern sich i.R. der Rückrechnung der unterjährigen Warenbestände ein vernünftiges (positives) Repräsentativ für den normalen Füllungszustand des Lagers ergibt, werden in einem zweiten Diagramm die (rechnerisch) **relativen Bestandsveränderungen** für die einzelnen Monate als

weitere Veranschaulichung abgebildet. Durch diese zusätzliche Perspektive auf die Warenflussdaten tritt im Beispiel eines größeren Gastronomiebetriebs der **fehlende Wareneinkauf im Dezember der ersten beiden Jahre** auf. Die Nachverfolgung dieser Regelstörungen nach der SRP-Zeitreihensystematik (s. [Pkt. 4.2.2.](#)) führt zu hohen Rückvergütungen. Dass diese **für das letzte Jahr nicht in der Rückrechnung zu erkennen** sind, liegt an der Funktion des Warenendbestands²⁵⁰ als Fixparameter für die retrograde Berechnung.



Die Nachverfolgung dieser Regelstörungen nach der SRP-Zeitreihensystematik (s. [Pkt. 4.2.2.](#)) führt zu hohen Rückvergütungen. Dass diese **für das letzte Jahr nicht in der Rückrechnung zu erkennen** sind, liegt an der Funktion des Warenendbestands²⁵⁰ als Fixparameter für die retrograde Berechnung.

4.2.9. Interaktive Optionen zur verbesserten Sachaufklärung

Ein fundamentaler Vorteil der digitalen Außenprüfung ist die Möglichkeit, neue Erkenntnisse jederzeit reversibel in die (Gesamtbild-)Prüfung einzupflegen und deren Auswirkungen auf die Ergebnisse umgehend zu simulieren. Dazu beinhalten die SRP-Vorlagen diverse Funktionen, um die Prüfungsdaten nachvollziehbar an neue Informationen anzupassen. Ein Plausibilitätsurteil i.S. des § 158 AO kann somit unter Berücksichtigung aller bekannten Fakten vorgenommen werden.

Während die interaktiven Korrekturoptionen im Bereich der musterbasierten Analysen der SRP-Vorlage Stochastik wegen der speziellen Prüfsystematik nicht in einem gesonderten Kapitel zusammengefasst sondern zu den jeweiligen Einzeluntersuchungen bzw. Untersuchungsschritten beschrieben werden (s. [Pkt. 4.3.1.](#) bis [Pkt. 4.3.4.](#)), gibt dieser Pkt. einen

²⁵⁰ Der Warenendbestand ergibt sich aus den Eingaben „Anfangsbestand“ (Tabellenblatt „Bestandsentwicklung“) sowie „Bestandsveränderungen“ (Tabellenblatt „Jahreszahlen“).

Überblick über die **Anpassungsoptionen der SRP-Vorlage BWA**. Folgende Funktionen werden dazu in den nächsten Abschnitten beschrieben:

- Einzelkorrektur von Datensätzen durch Ausblenden, Umdatierung oder Wertänderung,
- Einschränkung des Betrachtungszeitraums,
- Verschieben der ersten Datenreihe um eine feste Anzahl Tage,
- Bereinigung einseitiger (konjunktureller) Preiseinflüsse sowie
- Niveaueinpassung eines Grafen zur Verbesserung der Vergleichbarkeit beim Zeitreihenvergleich.

Mit Hilfe dieser Möglichkeiten lässt sich die gemeinsame Pflicht zur Sachverhaltsaufklärung²⁵¹ konkret und unter deutlich reduziertem Aufwand umsetzen. In der Konsequenz können Gesamtbildbeurteilungen zur Schlüssigkeit von Besteuerungsgrundlagen gem. § 158 AO qualifizierter getroffen und ggf. die Substanzlosigkeit der Mitwirkung von Stpfl.²⁵² eindeutig herausgestellt werden.

Sämtliche interaktiven Beeinflussungen der untersuchten Besteuerungsdaten lassen sich mit den druckvorbereiteten SRP-Vorlagen nachvollziehbar dokumentieren. So werden vorgenommene Korrekturen auf allen betroffenen Auswertungsblättern automatisch angezeigt, und es können in den Bemerkungsfeldern individuelle Kommentare ergänzt werden.

4.2.9.1. Einzelkorrektur von Datensätzen

Für Einzelkorrekturen (tlw.) fehlerhafter Datensätze stehen zu beiden Datenreihen in der

Einkauf				
		-564,38		
Datum	Betrag	Korrektur	Bemerkung	Einkauf korrigiert
857	853	4	5	856
29.08.2002	2.257,50	-2.257,50	Rg. 647 mit Bestands-einkauf Getränke	0,00
		564,38	Rg. 647 Verteilung nach Verbrauch	564,38
15.09.2002		564,38	Rg. 647 Verteilung nach Verbrauch	564,38
15.10.2002		564,38	Rg. 647 Verteilung nach Verbrauch	564,38
15.11.2002		564,38	Rg. 647 Verteilung nach Verbrauch	564,38
15.12.2002			Rg. 647 Verteilung nach Verbrauch	

SRP-Vorlage BWA jeweils zwei **Extraspalten** zur Verfügung, in denen **Wertberichtigungen und dazugehörige Erklärungen** eingetragen werden können (*Tipp: Zur Verbesserung der Nachvollziehbarkeit wird empfohlen, unterhalb*

der Datenreihe neu hinzugefügte Zeilen mit ihren Beträgen wie im Beispiel in die Korrektur-

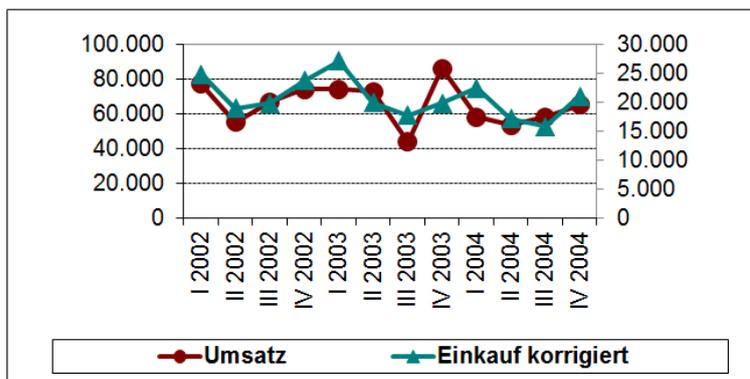
²⁵¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen“.

²⁵² Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

spalte zu schreiben.). Ein **Zählerkasten** protokolliert die Summe der bereits erfolgten Wert-
eintragungen, so dass schrittweise Umverteilungen erleichtert werden. Über die **Autofilter** zu
den Spalten lassen sich alle Anpassungen oder diejenigen zu einem besonderen Sachver-
halt gesondert ausgeben und dokumentieren.

Solange die vorgenommenen Einzelkorrekturen in ihrer Summe ungleich Null ausfallen,

erscheint zu jeder betroffenen
Auswertung automatisch ein
Hinweistext. Zusätzlich wird bei
insgesamt nicht neutralen Ein-
zelkorrekturen auf dem Tabel-
lenblatt „Schätzung“ mit einem
roten Rahmen um das Indi-
katorfeld „80 %-Quantil RGAS“
daran erinnert (s. [Pkt. 4.8.](#)). Soll
eine reine Umverteilung ohne



Auswertung nach Anpassung von Einzeldaten (siehe Vergleichsdaten)

Veränderung des Gesamtwerts der jeweiligen Datenreihe auf den Auswertungen erkennbar
sein, muss hierfür ein manueller Kommentar in die Bemerkungsfelder eingetragen werden.

4.2.9.2. Einschränkung des Betrachtungszeitraums

In einigen Fällen ist es bspw. zur Einschätzung einer beobachteten Verzerrung durch eine
außergewöhnliche Phase hilfreich, den Prü-
fungszeitraum zügig einschränken zu können.
Hierfür lassen sich auf dem Tabellenblatt „Ver-
gleichsdaten“ oben rechts auch **Teilzeiträume** frei eintragen, auf die dann alle Analysen
begrenzt werden.

Beginn Prüfungszeitraum:	01.01.2002
Ende Prüfungszeitraum:	15.06.2004
Abweichendes Wirtschaftsjahr:	

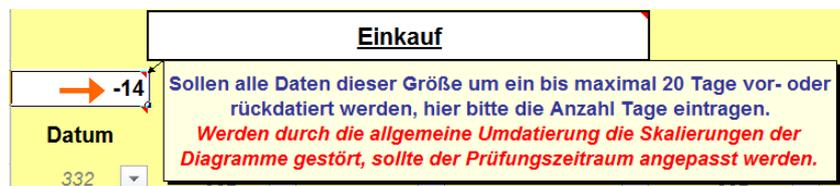
Ein gesonderter Hinweis hierzu ergeht auf den Auswertungsblättern nicht, weil den je-
weiligen Diagrammen i.d.R. die Unvollständigkeit des Prüfungszeitraums zu entnehmen ist.
Soll eine zeitliche Einschränkung dennoch zu den Analysen aufgeführt werden, muss die
Eingabe in den Bemerkungsfeldern per Hand erfolgen. Im Fall der Festlegung einer Schät-
zungshöhe mit der SRP-Vorlage BWA – *nach eingetretener Amtspflicht der Schätzung gem.
§ 162 Abs. 2 S. 2 AO* – wird durch einen roten Rahmen um das Indikatorfeld „80 %-Quantil
RGAS“ darauf hingewiesen, wenn Daten auf dem Tabellenblatt „Vergleichsdaten“ außerhalb
des eingetragenen Prüfungszeitraums liegen.

Im Hinblick auf die sachgerechte Bestimmung einer Schätzungshöhe unter wesentlich
abweichenden Bedingungen von Phasen des Prüfungszeitraums mit der Quantilsschätzung
(s. [Pkt. 4.8.](#)) – *nach eingetretener Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO* – ist

die Einschränkung des Analysezeitraums ebenfalls komfortabel, indem die SRP-Vorlage BWA mit den Daten kopiert und für die jeweiligen Teilzeiträume definiert wird.

4.2.9.3. Verschieben der ersten Datenreihe

Trotz der guten Stabilität des doppelskaliert größenangepassten Zeitreihenvergleichs gegenüber eher alltäglichen Störeinflüssen (s. den Vergleich der gängigen Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) wird bei Fällen mit Einnahmen-Überschuss-Rechnung nach § 4 Abs. 3 EStG regelmäßig als Erklärungsansatz für Regelstörungen der **Zeitversatz zwischen operativem Wareneinkauf bzw. -verbrauch und dem Erfassungszeitpunkt** der Bezahlung angeführt. Um diese – nach ständigen Prüfungserfahrungen häufig nicht für die entdeckten Auffälligkeiten ursächliche – Besonderheit interaktiv in die SRP-Zeitreihensystematik einbinden zu können, lässt sich die erste Datenreihe um bis zu 20 Tage²⁵³ komplett auf der Zeitachse nach hinten oder vorne verschieben. Beim exemplarischen Dönerimbiss, der in den (Quartals- und) Monatszeitreihen (s. [Pkt. 4.2.5.2.](#)) sowie der Rückrechnung des Warenbestands (s. [Pkt. 4.2.8.](#)) mit unschlüssigen Besteuerungsdaten zum Warenfluss aufgefallen ist, wird lt. Abb. eine simulierte **Rückdatierung aller Wareneinkäufe um 14 Tage** eingegeben, um den Einfluss der verzögerten Aufzeichnungs-



weise auf die Zeitreihenanalysen zu prüfen.

Durch die Vor- bzw. Rückdatierung aller Positi-

onen einer Datenreihe fallen i.d.R. einige aus dem eingegebenen Prüfungszeitraum heraus, wodurch die ersten oder letzten Periodensummen erheblich verzerrt werden können. Falls dadurch einzelne Zeitreihenanalysen infolge automatisch angepasster Y-Achsenkalierungen unbrauchbar scheinen, lässt sich dieser (negative) Effekt am schnellsten durch eine Einschränkung des Prüfungszeitraums auf dem Tabellenblatt „Vergleichsdaten“ aufheben.

Eine zeitliche Verschiebung der ersten Datenreihe hat auf jedem betroffenen Auswertungsblatt einen entsprechenden Hinweis unter Benennung der Anzahl und Richtung der eingetragenen Tage zur Folge.

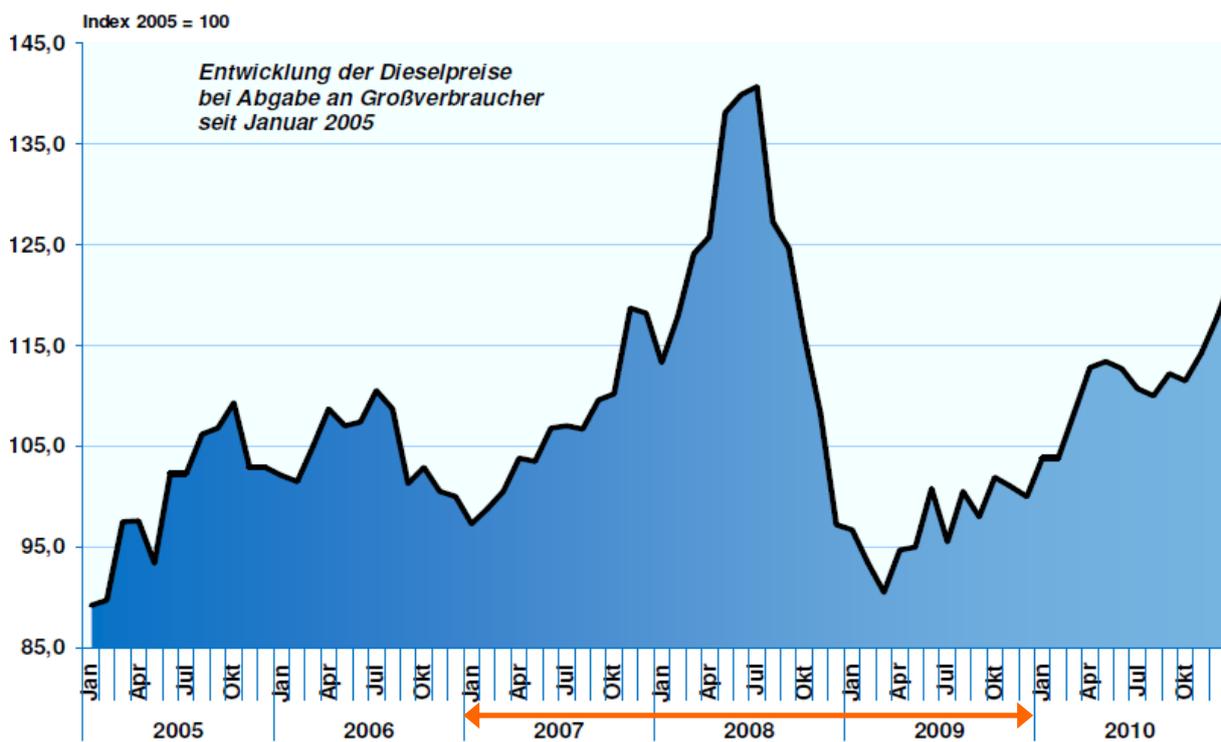
Weil die Vor- oder Rückdatierung einer kompletten Datenreihe spätestens i.R. einer Schätzung – **nach eingetretener Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO** –

²⁵³ Eine grundsätzliche Zahlungsferne aller Wareneinkäufe von mehr als 20 Tagen erscheint im Wirtschaftsleben, in dem Liquidität ein kostbarer Vorteil von Unternehmen ist, alleine schon implausibel i.S. des § 158 AO. Soll im Einzelfall dennoch eine weitreichendere Verschiebung einer Datenreihe simuliert werden, müssen die jeweiligen Daten entweder vor dem Hineinkopieren in die SRP-Vorlage BWA oder auf dem Tabellenblatt „eigene Notizen“ entsprechend vollständig vor- oder rückdatiert werden.

zurückgenommen werden sollte, erinnert das Indikatorfeld „80 %-Quantil RGAS“ mit roter Schrift an die bestehende Aktivierung dieser interaktiven Funktion (s. [Pkt. 4.8.](#)).

4.2.9.4. Beseitigung einseitiger (konjunktureller) Preiseinflüsse

Ein wichtiger Störfaktor für Zeitreihenergebnisse können einseitige konjunkturelle Einflüsse sein, auf die in Unternehmen nur bedingt reagiert werden kann. Dazu gehören kurzfristige starke Preiserhöhungen im Einkauf (s. *das Beispiel unter [Pkt. 4.2.5.1.](#)*), die nicht an die Kunden weitergegeben werden sollen, sowie (gesetzliche) Preisbindungen. Obwohl der doppel-skaliert größenangepasste Zeitreihenvergleich stabil gegenüber Störeinflüssen ist (s. *den Vergleich gängiger Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)*), bietet die SRP-Vorlage BWA mit der **Konjunkturbereinigung** eine Funktion, um derartige einseitige Preiseffekte zu egalisieren²⁵⁴. Im Ergebnis würde durch die (vollständige) Umrechnung aller Werte beider Datenreihen mittels Indizes²⁵⁵ auf ein einheitliches Preisniveau aus einer Betrags- eine Mengenverproben werden.

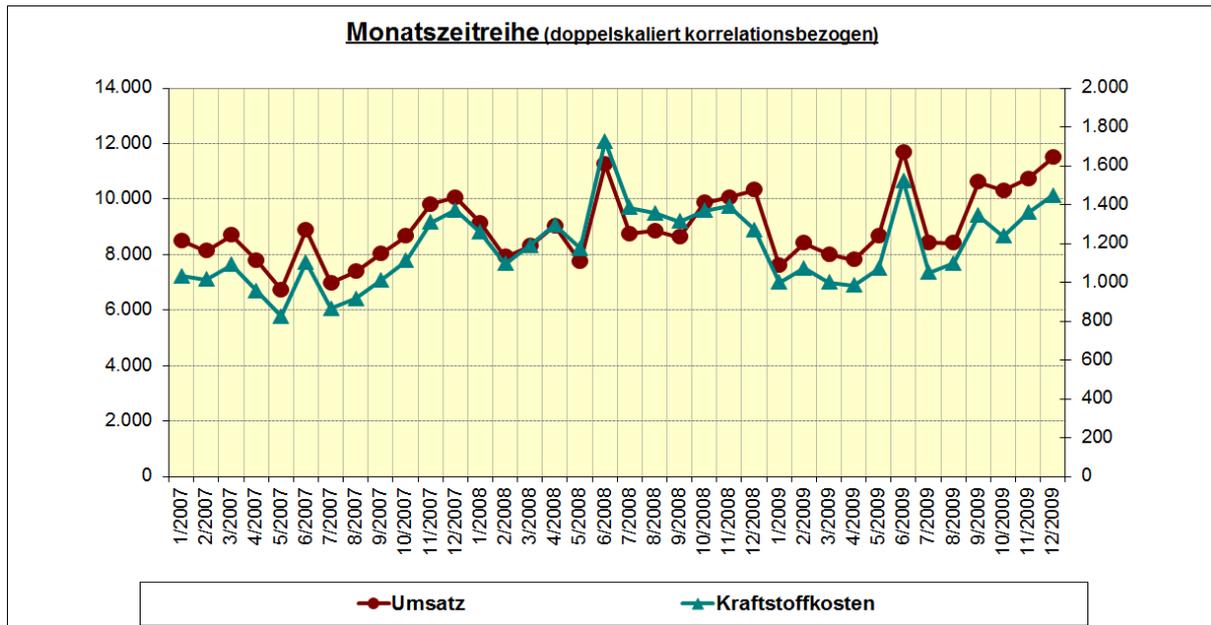


²⁵⁴ Die Funktion der – in der Volkswirtschaftslehre üblichen – Bereinigung von Preiseffekten durch Indizes ist vom BFH (Beschl. v. 12.7.2017 – X B 16/17, Rz. 128) bisher nicht ausreichend gewürdigt worden: „Das FA hat hierzu in der Beschwerdeerwiderung ausgeführt, mit der ‚Konjunkturanpassung‘ hätten die Ergebnisse der Jahre 2009 und 2010 an das Preisniveau des Jahres 2008 angepasst werden sollen, um die Vergleichbarkeit der Schätzung herzustellen. Indes soll eine Schätzung nicht vergleichbar, sondern möglichst realitätsnah sein. Wenn im Prüfungszeitraum eine erhebliche Preiserhöhung stattgefunden hat, dann muss die Schätzung dies berücksichtigen, nicht aber die – nicht miteinander vergleichbaren – RAS der Jahre vor und nach der Preiserhöhung vergleichbar machen.“

²⁵⁵ Preisindizes rechnen Werte von „nominal“ auf „real“ um, wonach sie denselben Preisverhältnissen unterliegen und vergleichbar werden.

In der Taxibranche haben einseitige Konjunkturreffekte eine besondere Praxisbedeutung, weil den veränderlichen Kraftstoffpreisen als wesentlichem Ausgabenfaktor festgesetzte Tarife gegenüberstehen. Durch die regionalen Preisbindungen für Taxifahrten kann nur bedingt auf Marktschwankungen reagiert werden. So hat für einen Beispielsfall der Dieselpreis während des Prüfungszeitraums „2007 bis 2009“ lt. Abb. oberhalb erheblich geschwankt²⁵⁶:

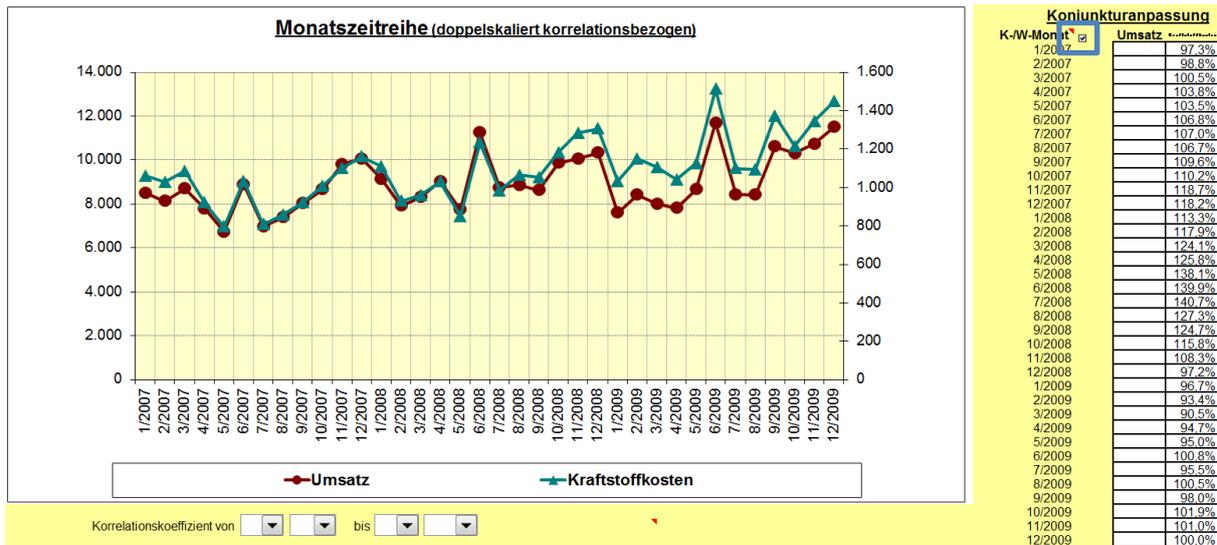
Dennoch lässt der doppelskaliert größenangepasste Zeitreihenvergleich grundsätzlich eine klare Abhängigkeit des Dieselkaufs zum Absatz erkennen:



Auf (fast) jede Veränderung der Nachfrage reagiert der Kraftstoffeinkauf gleichgerichtet und mit ähnlicher Intensität. Auffällig ist hingegen die Niveauerhöhung der Ausgaben im mittleren Jahr, welches genau die Phase des dauerhaft starken Anstiegs der Dieselpreise markiert.

Im Tabellenbereich „Konjunkturanpassung“ neben dem Diagramm lassen sich in die SRP-Vorlage BWA die monatlichen Preisindizes für eine oder beide Datenreihen einpflegen, um anschließend per **Anhakfeld** (s. die Abb. unterhalb) zwischen der konjunkturbereinigten Realversion des Zeitreihenvergleichs und der nominalen hin- und herschalten zu können. Auch mit diesem Extrembeispiel für Zeitreihenanalysen wird die Robustheit des doppelskaliert größenangepassten Zeitreihenvergleichs (s.a. den Vergleich von gängigen Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter **Pkt. 4.2.**) bestätigt, indem eine deutlich geringfügigere Veränderung resultiert, als aufgrund der eklatanten einseitigen Preisschwankungen zu erwarten war.

²⁵⁶ Quelle: Internetseite des Bundesverbands Güterkraftverkehr Logistik und Entsorgung (BGL) e.V., nachzulesen unter: <https://www.bgl-ev.de/images/downloads/dieselpreisinformation.pdf>. Auch wenn die tatsächlichen Preisschwankungen für das Beispielsunternehmen regional und als Endabnehmer von dieser Internetrecherche abweichen können, verdeutlicht die Gegenüberstellung der Zeitreihenvergleiche „nominal“ und „real“ den begrenzten Einfluss von Preiseffekten auf robuste Zeitreihenvarianten.



Neben der besseren Deckungsgleichheit beider Grafen in der ersten Hälfte trotz des starken Preisanstiegs des Kraftstoffs lässt sich eindeutig ablesen, dass ab Mitte 2008 der preisbereinigte Dieselaufwand im Verhältnis angewachsen ist. Diese – vorher unsichtbare – Regelstörung kann nun auf mögliche Verursacher wie z.B. die Anschaffung neuer Fahrzeuge mit einem höheren Verbrauch untersucht und hinterfragt werden. Ergeben sich keine überprüfbar betrieblichen Erklärungen, besteht Grund zu zweifeln, „dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt“²⁵⁷.

Falls die Konjunkturbereinigung aktiviert ist, wird dies zu den plausibilitätsrelevanten Auswertungen i.S. der SRP angezeigt, wobei eine Anwendung der Preisindizes vom Tabellenblatt „Monatszeitreihen“ auf die „Quartalsverhältnisse“ über das dortige Anhakfeld oben links aktiv angesteuert werden muss. Auf die Tabellenblätter „Gleitschlitzenzeitreihen“, „Monatsverhältnisse“, „Wochentagswerte“ und „Bestandsentwicklung“ wird die Konjunkturanpassung nicht übernommen, weil diese Analyseschritte i.R. der SRP nicht einzeln für eine Plausibilitätsbewertung i.S. des § 158 AO herangezogen werden.

Auf dem Tabellenblatt „Schätzung“ – *für den Fall der eingetretenen Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO* – wird an die Datenumrechnung zur Konjunkturbereinigung durch rote Schrift im Indikatorfeld „80 %-Quantil RGAS“ erinnert (s. [Pkt. 4.8.](#)).

4.2.9.5. Niveaueinpassung eines Grafen

In seltenen Fällen führt die automatische Größenanpassung durch eine zweite Y-Achse bei doppelskalierten Zeitreihenvergleichen nicht dazu, dass die beiden Grafen überwiegend aufeinander liegen und ihre Abhängigkeit somit besser beurteilt werden kann. Für diese Ausnahmen verfügt der einzelskalierte Zeitreihenvergleich auf dem Tabellenblatt „Monats-

²⁵⁷ Anwendungserlass zu § 158 AO.

zeitreihen“ über eine Umschaltfunktion auf quasi-prozentuale²⁵⁸ Darstellung mit der zusätzlichen Möglichkeit, die **erste Datenreihe im Diagrammniveau anzupassen**, bis eine Bewertung der Gesamtschlüssigkeit i.S. des § 158 AO seriös möglich wird (s. *das Beispiel unter [Pkt. 4.2.5.2.](#)*).

Da die Niveauanpassung der ersten Datenreihe beim einzelskalierten Zeitreihenvergleich keine Auswirkung auf die Größen anderer Auswertungen hat, ist ein warnender Hinweis auf anderen Analyseblättern bzw. i.R. einer Schätzung – *unter eingetretener Amtspflicht der Schätzung gem. § 162 Abs. 2 S. 2 AO* – nicht erforderlich.

4.2.10. Kosten- und weitere Zeitreihen

Neben den vielfältigen **Einsatzmöglichkeiten** zur Erstorientierung in einem Prüfungsfall sowie zur Bearbeitung von komplexen Einzelprüffeldern (s. [Pkt. 5.1.](#)) ergeben sich für Zeitreihenanalysen auch i.R. von Plausibilitätsprüfungen gem. § 158 AO praxisrelevante Ansätze über die Verbildlichung des dokumentierten Warenflusses hinaus.

Die Fokussierung der Finanzrechtsprechung auf die gegenüber Störeinflüssen anfälligen Aufschlagsatzzeitreihen²⁵⁹ (s. *den Vergleich der gängigsten Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)*) mit daraufhin festgelegten Anwendungsbeschränkungen²⁶⁰ hat jedoch den Eindruck eines schmalen Einsatzbereichs für die Zeitreihenanalytik verfestigt. Dagegen hat *Wolenski*²⁶¹ mit dem Beispiel eines bilanzierenden Unternehmens, welches Fremdleistungen auf Abruf einkauft und weiterveräußert, plastisch demonstriert, dass auch in anders gelagerten Fällen offensichtlich werden kann, dass „es nach Verprobung usw. unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen

²⁵⁸ Die Monatswerte werden dann als Anteile der jeweiligen Jahressummen angegeben.

²⁵⁹ In den ausführlichen Zeitreihenbeschreibungen der Finanzrechtsprechung dominiert mit der 10-Wochen-Gleit-schlittenzzeitreihe eine Aufschlagsatzvariante wie z.B. in: BFH v. 25.3.2015, BStBl. II 2015, 743, FG Münster v. 19.8.2004 – 8 V 3055/04, FG Düsseldorf v. 15.2.2007 – 16 V 4691/06 oder FG Köln v. 27.1.2009 – 6 K 3954/07. In Anbetracht der Verbreitung der SRP über drei Viertel der Bundesländer und der Tatsache, dass die Gleitschlittenmodelle i.R. der SRP nicht zur Plausibilitätsbeurteilung herangezogen werden (s. [Pkt. 4.2.6.](#)), kann die extrem kleine Auswahl an Finanzgerichtsverfahren nicht als aussagekräftiges Repräsentativ für die Methodenverbreitung angesehen werden. Das FG Hamburg (v. 31.10.2016 – 2 V 202/16, Rz. 48) beschreibt entsprechend konkret die Andersartigkeit des doppelskaliert größenangepassten Zeitreihenvergleichs als zentrale SRP-Variante für die Schlüssigkeitsbeurteilung: „Mit Hilfe der doppelskalierten Monatszeitreihe hat die Betriebsprüfung die Daten des Wareneinkaufs den Umsatzerlösen in zwei Einzelgrafiken gegenüber gestellt. Bei dieser Methode tritt das zahlenmäßige Verhältnis zu Gunsten der Abhängigkeit voneinander in den Hintergrund. Wesentlich ist die Reaktion der Faktoren aufeinander. Aus kaufmännischer Sicht wird in einem Unternehmen auf jede Veränderung der Nachfragesituation (annähernd) adäquat mit einer Erhöhung bzw. Absenkung der Einkaufssumme reagiert“.

²⁶⁰ BFH v. 25.3.2015, BStBl. II 2015, 743, Leitsatz 1: „Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist.“

²⁶¹ *Wolenski*, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“, Stbg 6/2016, 268, 271.

Verhältnissen übereinstimmt²⁶²: Es lassen sich in zwei Jahren zum Jahresende hin Fremdleistungseinkäufe (fast) in Millionenhöhe ablesen, obwohl keine – auch nur annähernd – gleichwertigen Verkäufe folgen. Da zudem jeweils keine „teilmfertigen Arbeiten“ bilanziert worden sind, verbleibt nur die logische Schlussfolgerung, dass die Erlöserfassung unvollständig sein muss.

Gerade weil „[e]lektronische Kassensysteme [...] durch Umprogrammierung in nahezu beliebiger Weise manipulierbar [sind und] von derartigen Manipulationsmöglichkeiten [...] Teile der betrieblichen Praxis nach dem Erkenntnisstand des Senats durchaus Gebrauch [machen]“²⁶³ setzt eine effektive Prüfung von Besteuerungsgrundlagen auch Ansätze voraus, die durch digitale Hilfsmittel [s.a. den Abschnitt „raffinierte (digitale) Manipulationswerkzeuge“ unter [Pkt. 2.1.](#)] deutlich schwieriger systematisch zu unterwandern sind wie bspw. **Kostenzeitreihen**. Der Plausibilitätsansatz des § 158 AO setzt an derartigen Zeitreihenvergleichen mit derselben Schlüssigkeitsfrage an, ob sich die Unternehmensleitenden mit dem Aufwand für variable Kosten an der Ertragssituation orientieren, um den Gewinn zu maximieren.

Mit seiner Stabilität gegenüber alltäglichen Störeinflüssen (s. den Vergleich der gängigsten Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) ist der doppelskaliert größenangepasste Zeitreihenvergleich auch für einen kostenbezogenen Einsatz gut geeignet. Bei der Auswahl der zur Gegenüberstellung herangezogenen Aufwandsfaktoren ist allerdings darauf zu achten, dass viele Kostenpositionen (zeitweise) durch Alternativen ersetzt werden können (Substitution). Deshalb bietet sich für Kostenzeitreihen besonders die variable Arbeitsweise mittels Pivot-Tabelle auf dem Tabellenblatt „eigene Notizen“ (s. den Abschnitt „fortgeschrittene Arbeitsweise der SRP-Vorlagen“ unter [Pkt. 4.](#)) an. Über die [Berichtsfiler](#) bzw. Datenschnitte²⁶⁴ können die ausgewählten Kostenkonten nun jederzeit angepasst werden.

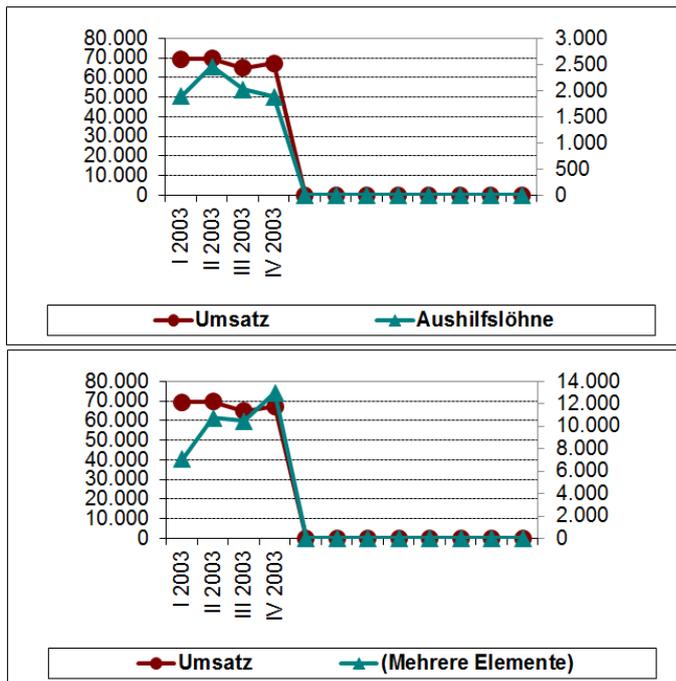
Kontonummer	(Alle)
Kontobezeichnung	Aushilfslöhne
Zeilenbeschriftungen	Summe von SoHa
04.01.2003	148,8
06.01.2003	489,1
04.02.2003	758,8
04.03.2003	500,8
01.04.2003	258
03.04.2003	527,7

So offenbart der nachfolgend abgebildete Zeitreihenvergleich der Aushilfslöhne zum Umsatz im Beispielsfall „Dönerimbiss“, welcher bereits durch ungeschlüssige Aufzeichnungen zum Warenfluss aufgefallen ist (s. [Pkt. 4.2.5.2.](#) sowie [Pkt. 4.2.8.](#)), dass bereits auf der groben Übersichtsebene der Quartalssummen (s. den SRP-Zeitreibenduktus „vom Groben zu

²⁶² Anwendungserlass zu § 158 AO.

²⁶³ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

²⁶⁴ Während die älteren Berichtsfiler voneinander unabhängig sind, erkennen Datenschnitte automatisch den Zusammenhang. So zeigt der Berichtsfiler „Kontonummer“ im Beispielsfall keine Einschränkung an, obwohl über die Kontobezeichnung nur die Aushilfslöhne ausgewählt wurden. Bei Datenschnitten würde die Auswahl unter „Kontobezeichnung“ auf die Anzeige zur Kontonummer durchwirken.



den Details“ unter [Pkt. 4.2.2.](#)) kein erkennbarer Zusammenhang zwischen der konstanteren Absatzentwicklung und dem schwankenden variablen Kostenfaktor besteht.

Zur Absicherung, ob diese Regelstörung auf eine Umstellung im Bereich der Festangestellten zurückzuführen sein könnte, wird der Berichtsfiler mit der Auswahl der Vergleichskosten auf das Konto „Löhne und Gehälter“ ausgeweitet. Der aktualisierte doppelskalierte Zeitreihenvergleich klärt jedoch umgehend auf, dass diese

mögliche Strukturveränderung²⁶⁵ im Betrieb nicht ursächlich für das unschlussige Verhältnis der beiden Datenreihen ist.

Mit den Tabellenblättern der SRP-Vorlage BWA und einem entsprechenden Hinweis auf den erweiterten kostenbezogenen Zeitreihenvergleich im Bemerkungsfeld lässt sich die ergänzende Untersuchung zur Schlüssigkeit der Bestimmungsaufzeichnungen i.S. des § 158 AO nachvollziehbar dokumentieren. Im Fall konkreter Einlassungen²⁶⁶ können die Argumente jederzeit interaktiv (s. [Pkt. 4.2.9.](#)) eingepflegt und ihre Auswirkungen auf die Zeitergebnisergebnisse überprüft werden. Auf diese Weise lässt sich die Mitwirkungsqualität²⁶⁷ als Unterscheidung von wichtigen neuen Informationen für das Gesamtbild der Bestimmungsaufzeichnungen einerseits und Schutzbehauptungen andererseits feststellen.

Eine weitere erfolgversprechende Option zur verbesserten Aufdeckung systematischer Einnahmenunterdrückungen mit Hilfe digitaler Aufzeichnungsweisen [s.a. den Abschnitt „raffinierte (digitale) Manipulationswerkzeuge“ unter [Pkt. 2.1.](#)] ergibt sich aus **Zeitreihenver-**

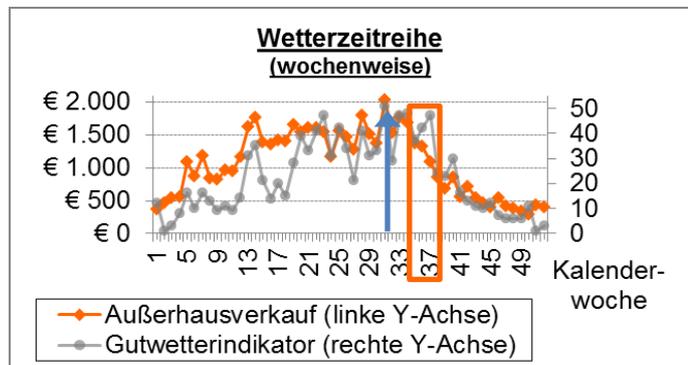
²⁶⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, 1. Leitsatz: „Es darf zudem im maßgebenden Zeitraum nicht zu solchen Änderungen in der Betriebsstruktur gekommen sein, die – nicht anderweitig behebbar – wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringen.“

²⁶⁶ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

²⁶⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

gleichen mit externen Daten. Auf der einen Seite lassen sich in einem kompletten Betriebsverwaltungssystem Unternehmensdaten zusammenhängend unauffällig manipulieren. Im Verhältnis zu externen Einflussfaktoren ist dies jedoch nur möglich, wenn diese Informationen in das System eingespeist wurden. Auf der anderen Seite sind heutzutage viele Daten zu Faktoren, die den Erfolg von Unternehmen und damit die Handlungslogik beeinflussen, frei im Internet zugänglich und können für Plausibilitätsuntersuchungen herangezogen werden. Ähnlich der interaktiven Berichtigung von einseitigen Preiseffekten bei einem Taxiunternehmen (s. [Pkt. 4.2.9.4.](#)) muss im Fall einer Beanstandung zwar die Anwendbarkeit der externen Größen auf den einzelnen Betrieb überprüft werden. Der berechtigte Ruf nach wirksameren Prüfungsmethoden gegen moderne Manipulationen²⁶⁸ unterstützt aber mögliche Ansätze, die denklogisch eine höhere Aufdeckungswahrscheinlichkeit erreichen.

In diesem Beispiel einer Wetterzeitreihe werden die externen Wetterdaten zur Schlüssigskeitsprüfung der Erlöse eines Eiscafés eingesetzt. Dazu sind die Außerhausverkäufe wochenweise aufsummiert worden, um die Informationsfülle der Auswertung zu begrenzen und dennoch die erwartet schnellen Nachfragereaktionen auf Veränderungen des Wetters erkennen zu können.



Die absatzsteigernden Wetterparameter sind zur besseren Übersichtlichkeit in einem Gutwetterindikator²⁶⁹ verarbeitet und ebenfalls als Wochendurchschnitt berechnet worden. Der Ergebnisgrafik ist einerseits zu entnehmen, dass die dauerhafte Wetterverbesserung vor der 21. Kalenderwoche keine adäquate Absatzreaktion ausgelöst hat²⁷⁰, obwohl z.B. die 31. Kalenderwoche offenlegt, dass weiteres Steigerungspotential vorhanden war. Als punktuelle Implausibilität wird die zweiwöchige Gutwetterperiode um die 37. Kalenderwoche nicht von einer Zunahme des Außerhausverkaufs beantwortet.

²⁶⁸ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 54: „Denn auch umgekehrt verschafft sich ein Teil der Steuerpflichtigen durch ausgefeilt geplante Doppelverkürzungen und/oder den Einsatz von Manipulationssoftware, die die Aufdeckungsmöglichkeiten herkömmlicher Prüfungsmethoden minimieren, technische Vorteile gegenüber der Finanzverwaltung [...] Hierauf darf und muss die Außenprüfung im Interesse der Wahrung der Gesetzmäßigkeit und Gleichmäßigkeit der Besteuerung (Art. 3 Abs. 1, Art. 20 Abs. 3 GG) reagieren, und zwar auch im Wege der Entwicklung und Anwendung neuartiger Prüfungsmethoden.“

²⁶⁹ Um den erwartet stärkeren Einfluss einer zusätzlichen Sonnenstunde gegenüber einer Temperaturerhöhung um 1° C. auf die Wettergüte zu beachten, wird ein Gutwetterindikator mit folgender Formel angesetzt: $Sonnenstunden \times 3 + Durchschnittstemperatur$. Die Prüfungserfahrungen mit dem Ansatz haben gezeigt, dass offensichtliche Unschlüssigkeiten auch bei einer anderweitigen Gewichtung im Gutwetterindikator erhalten bleiben.

²⁷⁰ Nach den „Fuzzy-Theorien“, die in der SRP-Zeitreihe-systematik berücksichtigt werden, wäre dieser Langzeittrend unter längeren Zeitreihenperioden besser zu erkennen (s. [Pkt. 4.2.2.](#)).

Für eine qualifizierte Beurteilung derartiger Auffälligkeiten i.S. der Plausibilität von Besteuerungsgrundlagen gem. § 158 AO ist die Aktualisierung des Zeitreihenvergleichs nach konkreten Einlassungen²⁷¹ mit neuen gesamtbildrelevanten Informationen jederzeit interaktiv (s. [Pkt. 4.2.9.](#)) möglich, so dass echte Mitwirkung von reinen Schutzbehauptungen (weitgehend) unterschieden werden kann²⁷².

4.2.11. Beweiskraft von (regelbasierten) Zeitreihenanalysen

Gesamtbildprüfungen (Verprobungen) sollen die Vergangenheit anhand von Unterlagen, Auskünften und Eindrücken wiederherstellen. Da Informationen zur Vergangenheit niemals vollständig und objektiv sind, ergibt sich grundsätzlich eine **Sachverhaltsbeschreibung mit Unsicherheiten**. Bei Gesamtbildprüfungen ist diese Unsicherheit besonders groß, weil aufgrund der Komplexität der entscheidungsrelevanten Vorgänge keine einfachen Ursache-Wirkungs-Beziehungen vorliegen können²⁷³.

Um dennoch in Gerichtsverfahren und Prüfungen Urteile bzw. ausreichend sichere Beurteilungen zu ermöglichen, werden an den juristischen Beweis geringere Anforderungen gestellt als an den naturwissenschaftlichen. Für den Bereich des Besteuerungsverfahrens sieht das Gesetz auf Seiten der Finanzverwaltung in **§ 158 AO (Verifikationsprinzip)** eine abermals abgesenkte erforderliche Aussagegewissheit vor (s. [Pkt. 4.6.](#)), denn andernfalls ließe sich die (annähernde) Gleichbehandlung aller Stpfl. gem. Art. 3 GG in einem Massenverfahren nicht umsetzen²⁷⁴.

Die vom BFH²⁷⁵ zur Schätzung – und der Argumentation folgend ebenso zur Verprobung – bevorzugten Methoden „Vermögenszuwachs-“, „Geldverkehrsrechnung“ und „Kalkulation“

²⁷¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweiszeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen. Den Stpfl. seinerseits verpflichten die Bedenken etc. zugleich, durch die Erfüllung seiner Mitwirkungspflichten (§ 90) zu eben dieser Sachverhaltsaufklärung beizutragen.“

²⁷² Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

²⁷³ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15: „Unser Dilemma im Umgang mit der Komplexität unserer Welt lässt sich darauf zurückführen, dass wir wohl darin ausgebildet wurden einfache logische Schlüsse zu ziehen und nahe liegende Ursache-Wirkungs-Beziehungen zu definieren. [...] Simple Ursache-Wirkungs-Beziehungen gibt es nur in der Theorie, nicht in der Wirklichkeit.“

²⁷⁴ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

²⁷⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 63 u. 64 : „Allein die Ergebnisse eines Zeitreihenvergleichs lassen aufgrund der dieser Verprobungsmethode innewohnenden methodenbedingten Unsicherheiten aber noch keinen sicheren Schluss auf das Vorliegen und den Umfang auch materieller Unrichtigkeiten der Buchführung

setzen bereits bei ordnungsmäßigen Besteuerungsunterlagen auf unvollständigen Ausgangsinformationen auf, denn bspw. private Geldzu- und -abflüsse bzw. der Warenverderb müssen nicht aufgezeichnet werden. Zudem ist das Risiko ordnungswidriger Besteuerungsunterlagen in den betreffenden Fällen mit überwiegend Bareinnahmen besonders hoch.²⁷⁶ Mindestens unter diesen Umständen kann das Modell „Schwarzeinnahmen zur Bezahlung von Schwarzlöhnen“ mit den Mitteln einer Betriebsprüfung häufig nicht ausgeschlossen werden. I.R. von Kalkulationen liegen oft keine vollständigen und vertrauenswürdigen Informationen zu Verkaufspreisen, Sonderverkaufspreisen, Anwendungsquoten von Sonderverkaufspreisen, Wareneinkäufen etc. vor. In der Konsequenz müssen der Geldverbrauch bzw. der Sollumsatz auf unsicheren Zahlen rekonstruiert werden, so dass – trotz scheinbar eindeutiger Zielgröße – regelmäßig eine beträchtliche Unsicherheit verbleibt.

Sorgfältige Ermittlungen der Finanzbehörde²⁷⁷ sollen die Unsicherheiten soweit möglich begrenzen. Nach dem Verifikationsprinzip haben aber auch die Mitwirkungsbereitschaft und -qualität der Stpfl. eine wesentliche Auswirkung auf das erforderliche Beweismaß für das Auslösen der Schätzungspflicht.²⁷⁸ Bei Zeitreihenanalysen bedeutet das sorgfältige Erarbeiten eines Gesamtbilds, eine nachvollziehbare und möglichst robuste²⁷⁹ Schlüssigkeitsfrage voranzustellen, eine geeignete Zeitreihenvariante (*s. den Vergleich der gängigsten Zeitreihenmodelle im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)*) für die Unter-

zu. In diesen Fällen sind andere Schätzungsmethoden, die auf betriebsinternen Daten aufbauen oder in anderer Weise die individuellen Verhältnisse des jeweiligen Steuerpflichtigen berücksichtigen (z.B. Vermögenszuwachs- oder Geldverkehrsrechnung, Aufschlagkalkulation) grundsätzlich vorrangig heranzuziehen.“

²⁷⁶ Teutemacher, „Handbuch zur Kassenführung – Praxishandbuch für die rechtssichere Umsetzung“, NWB Verlag, 2015, 2: „Betriebsprüfungen bei Betrieben, die überwiegend Bargeschäfte tätigen, führen in 95 % der Fälle zu dem Ergebnis, dass die Kassenführung verworfen wird, weil die vorgelegten Bücher bzw. Aufzeichnungen sowie Daten nicht den Ordnungsvorschriften entsprechen und damit die sachliche Richtigkeit in Zweifel gezogen wird“ mit Verweis auf Kalischke, „Betriebe bangen vor Steuerprüfern – Beamte werfen immer mehr Kassenbuchführungen“, Westfälische Nachrichten v. 15.3.2013.

²⁷⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Der Wortlaut des § 158 („Anlass zu beanstanden“) mag dafür sprechen, dass bereits bloße Beweisanzeichen, Zweifel oder Bedenken ausreichen, um die Vermutung auszuräumen oder zu entkräften. Es ist jedoch nicht so, dass die FinBeh. nur irgendwelche Beanstandungen auszusprechen hätte und der Stpfl. dann gehalten wäre zu beweisen, dass die formell richtigen Buchungen sachlich nicht falsch sind. Ein solcher Negativbeweis kann unmöglich verlangt werden. Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen.“

²⁷⁸ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

²⁷⁹ Je mehr betriebsinterne und externe Einflussfaktoren auf die Entwicklung oder Abhängigkeit von Betriebsgrößen einwirken können, umso schwieriger ist eine aussagekräftige Plausibilitätsbeurteilung. So muss bspw. i.R. des Ansatzes, aus Erlösen und einem variablen Kostenfaktor anhand eines Zeitreihenvergleichs die Handlungslogik in einem Unternehmen zu bewerten, stets auch die Substituierbarkeit (Austauschbarkeit) beachtet werden wie bei der Aushilfslöhne-Zeitreihe im Abschnitt „Kostenzeitreihen“ unter *Pkt. 4.2.10.*

suchung zu wählen sowie mögliche Störeinflüsse, die das Zeitreihenergebnis verzerrt haben können, zu beachten.

Im Hinblick auf die genannten Kriterien eines sorgfältig erarbeiteten – und damit aussagekräftigen – Gesamtbilds gründet sich die **SRP-Zeitreihensystematik** auf **folgende**

Analysegrundsätze und Annahmen:

- ☑ Die Schlüssigkeitserwartung, dass in Unternehmen auf veränderte Erfolgssituationen mit der Anpassung variabler Aufwandsfaktoren reagiert wird wie z.B. mit dem Warennachkauf auf den Absatz, stellt eine logische und direkte Handlungslogik dar. Infolge der direkten Gegenüberstellung von zwei betrieblichen Datenreihen durch ein geeignetes (s. *den nächsten Aufzählungspkt.*) Zeitreihenmodell in der SRP-Vorlage BWA werden keine Störfaktoren vermischt, sondern es können beobachtete Regelauffälligkeiten gezielt untersucht werden. Demgegenüber wird zu dem Schlüssigkeitsansatz eines konstanten Aufschlagsatzes, welcher in der Finanzrechtsprechung zurzeit noch dominiert²⁸⁰, in der Literatur vor vielfältigen möglichen Störeinflüssen gewarnt²⁸¹.
- ☑ Die zentrale Plausibilitäts-Zeitreihenvariante „doppelskaliert größenangepasster Zeitreihenvergleich“ der SRP-Zeitreihensystematik ist im Gegensatz zu anderen gängigen Zeitreihenarten sehr stabil gegenüber alltäglichen Sondereinflüssen, wie die Praxisbeispiele zur Einnahmen-Überschuss-Rechnung gem. § 4 Abs. 3 EStG (s. *Abschnitt „Methodenkenntnis“ unter Pkt. 4.2.*), zu lange haltbaren Waren (s. *Abschnitt „Robustheit des doppelskalierten Zeitreihenvergleichs“ unter Pkt. 4.2.*) und zu starken einseitigen Preiseffekten (s. *Pkt. 4.2.9.4.*) nachvollziehbar belegen: In all diesen für Zeitreihenanalysen schwierigen Fällen ist die Abhängigkeit der betrachteten Betriebsfaktoren auf monatlich aufsummierter Ebene problemlos ablesbar.
- ☑ Die SRP-Zeitreihensystematik verwendet mehrere Zeitreihenmodelle mit ihren jeweiligen Stärken in zielgerichteter Kombination, um – dem Prüfungsduktus „vom groben Überblick zu den Details“ folgend – den Zusammenhang der untersuchten Datenreihen zu verstehen (s. *Pkt. 4.2.2.*). Der Einstieg über die hochgradig verdichtete Quartalsebene stellt mit ihrer Übersichtlichkeit das Erkennen systemtischer Besonderheiten wie einer permanent verzögerten Erfassung sicher (s. *Pkt. 4.2.4.*). Das schrittweise Vertiefen über den doppelskalierten Zeitreihenvergleich (s. *Pkt. 4.2.5.1.*), die punktgenau eingrenzenden Gleitschlittenzeitreihen (s. *Pkt. 4.2.6.*) und die Einzelperspektiven zu beiden Daten-

²⁸⁰ Vgl. z.B.: FG Münster v. 19.8.2004 – 8 V 3055/04, FG Düsseldorf v. 15.2.2007 – 16 V 4691/06 oder FG Köln v. 27.1.2009 – 6 K 3954/07. Auch das BFH-Zeitreihenurteil (v. 25.3.2015, BStBl. II 2015, 743) bezieht sich mit der verwendeten Gleitschlittenzeitreihe auf eine Zeitreihenvariante des Aufschlagsatzes.

²⁸¹ Wiggen, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168, 171 f.: „2. Die Ermittlung des periodischen wirtschaftlichen Wareneinsatzes [...] 3. Das Fehlen von periodischen Inventurwerten [...] 4. Preisschwankungen und andere betriebliche Besonderheiten“.

reihen (s. [Pkt. 4.2.7.](#)) sorgen im Anschluss für größtmögliche Sicherheit im Erkennen *aller* gesamtbildrelevanten Sondersachverhalte.

- ☑ Schlussendlich können alle entdeckten bzw. von seiten der Stpfl. dargelegten Sonder Vorgänge, welche Zeitreihenergebnisse evtl. negativ beeinflussen können, nachvollziehbar interaktiv (s. [Pkt. 4.2.9.](#)) korrigiert werden, indem sich einzelne Datensätze (tlw.) löschen oder umdatieren, der Prüfungszeitraum (zeitweise) einschränken, die erste Datenreihe um eine feste Anzahl Tage verschieben, einseitige Preiseffekte über Indizes ausgleichen und/oder bei Versagen der automatischen Doppelskalierung eine Niveaueinpassung des einen Grafen vornehmen lassen.

Die hohe Aussagegewissheit einer Plausibilitätsprüfung i.S. des § 158 AO mit SRP ergibt sich nicht nur aus den einzelnen aufgelisteten Qualitätsmerkmalen, sondern aus der SRP-Zeitreihensystematik, die alle Vorteile durch einen zielführenden Prüfungsablauf verbindet. Darüber hinaus können die Ergebnisse einzelner Zeitreihenanalysen sowie das Gesamtbild zur Schlüssigkeit durch musterbasierte Häufigkeitsauswertungen *unter* [Pkt. 4.3.](#) weiter abgesichert werden.

In seiner **Rechtsprechung** hat der BFH den Einsatz von Zeitreihenanalysen erheblich eingeschränkt²⁸² und dies mit systemimmanenten Methodenschwächen begründet²⁸³. Dabei erwähnt der X. Senat zwar die „zahlreichen Varianten eines Zeitreihenvergleichs“²⁸⁴, nimmt in der allgemeinen Beurteilung aber keine Differenzierung vor²⁸⁵. Demgegenüber hat das FG Hamburg die Notwendigkeit der Unterscheidung zutreffend beschrieben: „Mit Hilfe der doppelskalierten Monatszeitreihe hat die Betriebsprüfung die Daten des Wareneinkaufs den Umsatzerlösen in zwei Einzelgrafiken gegenüber gestellt. Bei dieser Methode tritt das zahlenmäßige Verhältnis zu Gunsten der Abhängigkeit voneinander in den Hintergrund. Wesentlich ist die Reaktion der Faktoren aufeinander.“²⁸⁶ In den weiter ausführenden Beiträgen der urteilenden RiBFH ist anschließend z.T. der Bezug des Urteils zu der im betreffenden Fall ver-

²⁸² BFH v. 25.3.2015, BStBl. II 2015, 743, Leitsatz 1: „Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist.“

²⁸³ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 39 u. 63: „Auch bei einer formell und materiell ordnungsmäßigen Buchführung führt ein Zeitreihenvergleich denkllogisch immer zu einem Mehrergebnis gegenüber der Buchführung, da der höchste Rohgewinnaufschlagsatz aller Zehn-Wochen-Perioden des Jahres auf den Wareneinsatz für das gesamte Jahr angewendet wird. [...] Allein die Ergebnisse eines Zeitreihenvergleichs lassen aufgrund der dieser Verprobungsmethode innewohnenden methodenbedingten Unsicherheiten aber noch keinen sicheren Schluss auf das Vorliegen und den Umfang auch materieller Unrichtigkeiten der Buchführung zu.“

²⁸⁴ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 37.

²⁸⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 54: „Aus diesen Befunden – so sensibel sie unter den aufgezeigten Aspekten der folgerichtigen Methodik und des effektiven Rechtsschutzes sein mögen – folgt für den Senat auf der Grundlage des derzeitigen Erkenntnisstandes allerdings nicht das Ergebnis, die Methode des Zeitreihenvergleichs grundsätzlich zu verwerfen.“

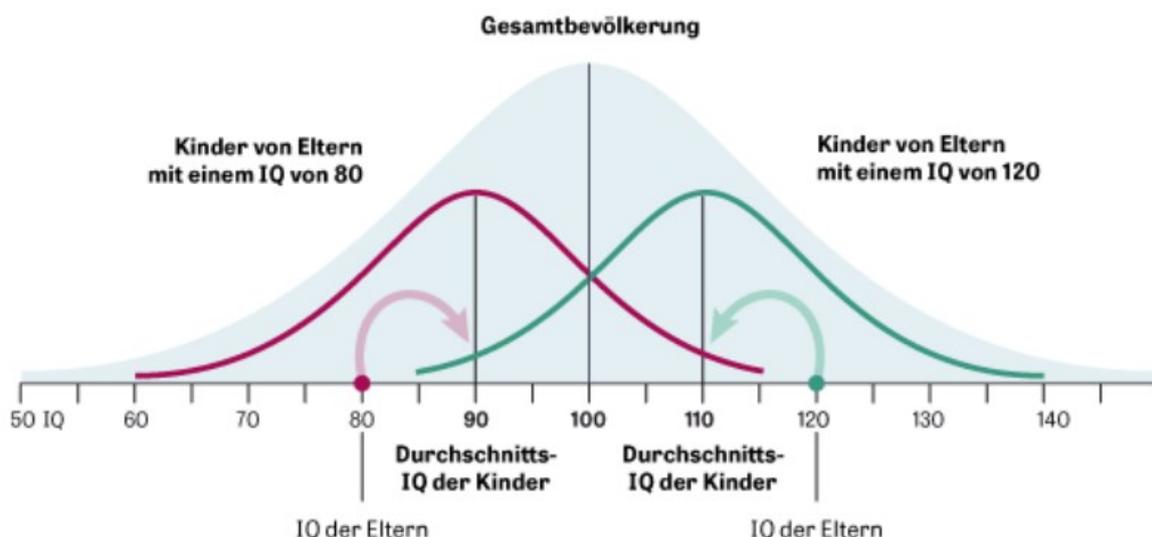
²⁸⁶ FG Hamburg v. 31.10.2016 – 2 V 202/16, Rz. 48.

wendeten 10-Wochen-Gleitschlitten-Aufschlagsatzzeitreihe bereits im Titel klargestellt worden.²⁸⁷

Da weder die vorangestellte Plausibilitätsfrage, noch die Zeitreihenvariante, die mehrperspektivische Systematik oder die interaktiven Optionen zur Einarbeitung neuer Erkenntnisse im Fall des BFH-Zeitreihenurteils mit der Arbeitsweise der SRP übereinstimmen, kann der einschränkende Richtspruch nicht auf die SRP übertragen werden.

4.3. Musterbasierte Analysen (Häufigkeitsauswertungen)

„Komplexität hat sehr viel mit Vernetzung zu tun, ja kommt erst durch Vernetzung zustande. Komplexe Vorgänge verlangen daher zu ihrem Verständnis ein Denken in Zusammenhängen, das sich an der Struktur organisierter Systeme und ihrer speziellen Dynamik orientiert. [...] Damit treffen wir auf einen der Kardinalfehler in der Beurteilung, was für die Erfassung von Komplexität wichtig ist. Denn hier geht es um das ‚Gesicht‘ der Wirklichkeit, um das Erkennen von Mustern“²⁸⁸. Mustererkennung als Analyseansatz setzt i.d.R. auf der **Auszählung von Informationen nach unterteilbaren**²⁸⁹ **Eigenschaften** auf. Diese Herangehensweise ist in der Wirtschaftsprüfung mit der ABC-Analyse seit langer Zeit zur Effizienzsteigerung bekannt, wenn die Daten eines bestimmten Bereichs nach ihrem Wertinhalt in „sehr wichtige“, „wichtige“ und „weniger wichtige“ differenziert werden. Mit der Variabilität und Schnelligkeit der digitalen Datenanalyse ist die Möglichkeit hinzugekommen, derartige Häufigkeitsauswertungen verbildlicht darzustellen.

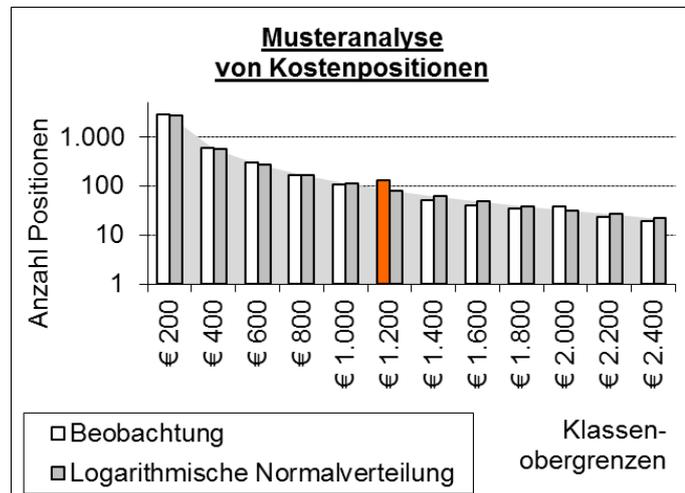


²⁸⁷ Nöcker (RiBFH), „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548.

²⁸⁸ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 16 u. 21.

²⁸⁹ Eigenschaften von Besteuerungsdaten, die sich zur Unterteilung eignen, sind z.B. die in Klassen eingeteilte Größe, der USt-Schlüssel, die Zahlungsweise etc.

Aus den Natur- und Wirtschaftswissenschaften ist bekannt, dass viele zusammengehörige Daten **gleichmäßige Muster der Häufigkeiten** wie z.B. die Normalverteilung in der Abb.²⁹⁰ oberhalb oder die Logarithmische Normalverteilung bilden²⁹¹. Deshalb haben sich tabellarische Häufigkeitsauswertungen zu visualisierten musterbasierten Analysen weiterentwickelt, mit denen eine zügige und intensive Durchdringung großer Datenmengen möglich ist. So beschreibt bspw. die nebenstehende musterbasierte Häufigkeitsauswertung mehrerer tausend Kostenpositionen eines Betriebs deren Größenzusammenhänge. In der Gegenüberstellung mit der Logarithmischen Normalverteilung, welche überwiegend typisch für Wirtschaftsdaten ist (s. [Pkt. 4.3.4.1.](#)), wird die Gleichmäßigkeit sofort sichtbar. Durch die



zielführende Darstellung mit logarithmischer²⁹² Y-Achse werden die Größenklassen mit wenigen Positionen betont, wodurch die **nicht mustergetreue Häufung in der Klasse „> € 1.000 bis € 1.200“** auffällt. Mit Hilfe von Filter- und Sortierfunktionen z.B. in IDEA® können die dort unerwartet geballt vorkommenden Kostenvorgänge unter geringem Aufwand herausgesucht und als betriebsbezogen sinnvolles Vertiefungsfeld abgeprüft werden.

Besonders die Kombination aus muster- und regelbasierten Analysen kann innerhalb kurzer Zeit eine hohe Transparenz trotz unprüfbar großer Datenmengen erzeugen. Deshalb werden die Funktionalitäten beider SRP-Teile (SRP-Vorlagen BWA und Stochastik) in der Prüfungssystematik zusammengeführt (s. *das Beispiel zur Konkretisierung eines auffälligen Zeitreihenvergleichs im Abschnitt „Prüfungsnetz“ unter [Pkt. 2.3.](#)*).

Natürliche gleichmäßige Muster in großen Datenmengen werden immer dann gestört, wenn Menschen durch eine der folgenden Aktionen in die Datenstruktur eingreifen:

 Manipulationen²⁹³,

²⁹⁰ Zeit online v. 17.6.2015: „Acht Antworten – Warum kluge Eltern nicht unbedingt kluge Kinder bekommen – und sieben weitere Tatsachen über die Vererbung von Intelligenz“, nachzulesen unter: <https://www.zeit.de/2015/23/intelligenz-vererbung-ig>.

²⁹¹ Vgl. z.B.: *Limpert/Stahel/Abbt*, „Lognormal distributions across the sciences: keys and clues“, BioScience Band 51, Nr. 5, 2001, 341, nachzulesen unter: https://de.wikipedia.org/wiki/Logarithmische_Normalverteilung.

²⁹² Eine logarithmische Skalierung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie in der Beispielsgrafik mit „10⁰ – 10¹ – 10² – 10³“ (entsprechend „1 – 10 – 100 – 1.000“) und erreicht dadurch eine Betonung im Bereich der kleinen Zahlen.

²⁹³ *Hill*, „The Difficulty of faking Data“, Chance 12/3/1999, 27.

-  Fehler bei der Erfassung von Daten z.B. durch Verschreiben, falsche Zuordnung, Schlüsselung etc. sowie
-  Gestaltungen i.d.R. in Form von Verträgen, wodurch die Eigenschaften der betroffenen Vorgänge bewusst festgelegt werden.

Auf der Basis dieser Erkenntnis sind **musterbasierte Prüfungsansätze zur Aufdeckung von Manipulationen** entwickelt worden. Nachdem diese in der Bekämpfung von Wirtschaftsverbrechen schon länger bekannt waren²⁹⁴, hat wohl Mark Nigrini²⁹⁵ den entscheidenden Anstoß gegeben, derartige Verfahren auch in der Wirtschaftsprüfung einzusetzen. Inzwischen sind musterbasierte Prüfungsmethoden zur Aufdeckung von unerwünschten menschlichen Eingriffen in großen Datenmengen sowohl in der Wirtschaftsprüfung²⁹⁶ als auch in der wissenschaftlichen Entwicklung²⁹⁷ vertreten. Da „[e]lektronische Kassensysteme [...] durch Umprogrammierung in nahezu beliebiger Weise manipulierbar [sind und] von derartigen Manipulationsmöglichkeiten [...] Teile der betrieblichen Praxis nach dem Erkenntnisstand des Senats durchaus Gebrauch [machen]“²⁹⁸ und dabei Steuerausfälle von knapp € 2 Mio. durch ein einzelnes Eiscafé entstehen können²⁹⁹ (s.a. [Pkt. 2.1.](#)), haben seit Einführung der digitalen Außenprüfung 2002 entsprechende Begriffe wie „Chi-Quadrat-Test“, „Benford-Gesetz“, „Strukturanalyse“ usw. auch in der Betriebsprüfung Verbreitung erlangt³⁰⁰.

²⁹⁴ Vgl. z.B.: Handelsblatt online v. 8.7.2012: „Volkswirte auf Verbrecherjagd: Ob illegaler Waffenhandel oder Steuerhinterziehung – Ökonomen stöbern immer häufiger Missetäter auf. Die forensische Ökonomie ist ein boomender Forschungszweig. Ermittlungsbehörden nutzen die Erkenntnisse der Forscher“, nachzulesen unter: <https://www.handelsblatt.com/politik/konjunktur/oekonomie/nachrichten/forensische-oekonomie-volkswirte-auf-verbrecherjagd/6838146.html> oder Handelsblatt online v. 30.11.2009: „Auf der Jagd nach Zahlen-Fälschern: Mit einem lange vergessenen mathematischen Gesetz wollen Ökonomen Kollegen überführen, die in wissenschaftlichen Studien Daten manipuliert haben. Denn längst nicht alle Wirtschaftswissenschaftler sind ehrlich – und die Anreize für Betrüger sind groß“, nachzulesen unter: <https://www.handelsblatt.com/politik/konjunktur/oekonomie/nachrichten/oekonomen-auf-der-jagd-nach-zahlen-faelschern/3315384.html>.

²⁹⁵ Nigrini, „The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies“, University of Cincinnati Ann Arbor Michigan, 1992; Internetseite von Mark Nigrini: <http://www.nigrini.com/>.

²⁹⁶ Vgl. z.B. den Einsatz der gleichmäßigen Verteilungen gem. Benford's Law (s. [Pkt. 4.3.2.3.1.](#)) als „general test“ i.R. der „in-house data analytics“ der Wirtschaftsprüfungsgesellschaft Ernst & Young, nachzulesen unter: <https://www.ey.com/UK/en/Services/Tax/VAT--GST-and-Other-Sales-Taxes/Tax---Indirect-Tax---ITTP---In-house-data-analytics>.

²⁹⁷ Vgl. z.B. die digitale Spurensuche mit modifizierter Benford-Verteilung (s. [Pkt. 4.3.2.3.1.](#)) am Fraunhofer-Institut für Sichere Informationstechnologie auf der Internetseite der Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V.: „Digitale Spurensuche – Wirtschaftsthemen Mai 2011“, nachzulesen unter: <https://www.fraunhofer.de/de/presse/presseinformationen/2011/mai/it-forensik-gegen-wirtschaftskriminalitaet.html>.

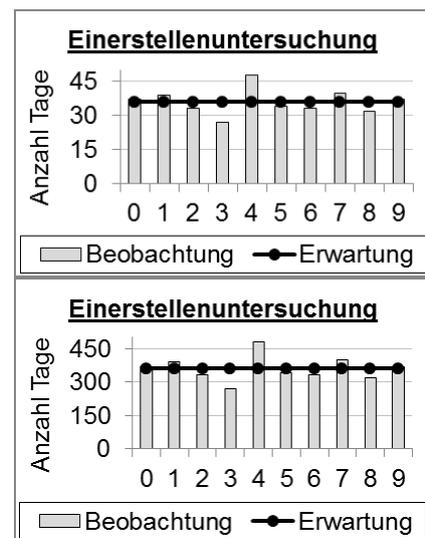
²⁹⁸ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

²⁹⁹ FG Rheinland-Pfalz v. 7.1.2015 – 5 V 2068/14; zitiert von: BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

³⁰⁰ Vgl. frühe Beiträge zu dem Thema wie z.B.: Sosna, „Einsatz statistischer Methoden zur Risikoanalyse, Recherche und Lokalisierung von Steuerausfällen“, StBp 2000, 41 u. 68, oder Huber, „Weiterentwickelte und neue Methoden der Überprüfung, Verprobung und Schätzung“, StBp 2002, 199, 233, 258 u. 293.

Eine wichtige Besonderheit zur Tendenz unbeeinflusster zusammengehöriger Datenbestände, gleichmäßige Muster einzunehmen, ist, dass die **Übereinstimmung mit dem Idealmuster** wie z.B. der Normalverteilung **von der Datenmenge abhängt**. Diese Verteilungseigenschaft ist aus der Stichprobenprüfung bekannt, indem die Genauigkeit, mit der Stichproben die Grundgesamtheit richtig wiedergeben, mit zunehmender Stichprobengröße (deutlich) besser wird (s. die Abb. „Abhängigkeit der Fehlerstreuung von der Stichprobengröße“ im Abschnitt „statistische Stichprobenverfahren“ unter [Pkt. 2.](#)).

An diesen beiden Beispielen zur Ziffernanalyse der Einerstelle aus Bareinnahmen (s. [Pkt. 4.3.2.2.](#)) lässt sich der Zusammenhang verdeutlichen: Beide Vergleiche zur ausgezählten Ein-Euro-Stelle aus manipulationsorientiert untersuchten Daten stimmen offensichtlich nicht mit der idealen Erwartung „Gleichverteilung aller Ziffern“ überein. Dabei zeigt die grafische Auswertung für beide Datenbestände ein (nahezu) identisches Abweichungsverhalten mit der deutlich zu häufigen „4“, der vergleichbar unterrepräsentierten „3“ und ansonsten kleineren Unterschieden zur erwarteten Gleichverteilung. An der Y-Achse lässt sich jedoch ablesen, dass im ersten Fall wenige Tageskasseneinnahmen bspw. aus einer abgekürzten Außenprüfung vorlagen, während im zweiten der zehnfache Datenumfang z.B. aus einer Durchsuchung verwendet werden konnte. Werden die Gegenüberstellungen der beiden Ziffernanalysen mit den Werkzeugen des mathematischen Spezialgebiets „Wahrscheinlichkeitsrechnung (Stochastik)“ beurteilt, gibt der χ^2 -Test (Chi-Quadrat-Test) für die obere Abb. ca. 53 % und für die untere 0 % Übereinstimmung aus. Aufgrund der wesentlich größeren Datenmenge wird im zweiten Beispiel eine erheblich bessere Mustertreue erwartet.



Aus der Gegenüberstellung der Beispiele wird deutlich, dass ein rein tabellarischer oder optischer Abgleich von beobachteten Mustern in Steuerungsdaten mit den jeweiligen Idealerwartungen zu falschen Schlussfolgerungen führen kann. Der mathematische Bereich der „**Wahrscheinlichkeitslehre (Stochastik)**“ bündelt die Theorie und Erfahrungen zum Streuverhalten von (zusammengehörigen) Daten und bietet komfortable Werkzeuge zur Übereinstimmungsprüfung (Wahrscheinlichkeitstests) an. Da Wahrscheinlichkeitstests durch die Kombinatorik³⁰¹ theoretisch reproduziert werden können und die beste Beschreibung

Aus der Gegenüberstellung der Beispiele wird deutlich, dass ein rein tabellarischer oder optischer Abgleich von beobachteten Mustern in Steuerungsdaten mit den jeweiligen Idealerwartungen zu falschen Schlussfolgerungen führen kann. Der mathematische Bereich der „**Wahrscheinlichkeitslehre (Stochastik)**“ bündelt die Theorie und Erfahrungen zum Streuverhalten von (zusammengehörigen) Daten und bietet komfortable Werkzeuge zur Übereinstimmungsprüfung (Wahrscheinlichkeitstests) an. Da Wahrscheinlichkeitstests durch die Kombinatorik³⁰¹ theoretisch reproduziert werden können und die beste Beschreibung

³⁰¹ Die (mathematische) Kombinatorik beschäftigt sich u.a. mit den möglichen Ergebnisvariationen (Permutationen) in diskreten – im Gegensatz zu kontinuierlichen – Datenverteilungen wie bspw. mit den Wahrscheinlichkeitsverhältnissen beim Werfen einer Münze. Durch das Aufstellen eines Ergebnisbaums aus den denkbaren Resultaten (bei zwei Münzwürfen mit den Seiten „A“ und „B“: „AA“, „AB“, „BA“ und „BB“) können die Wahrscheinlichkeiten für Einzelresultate (z.B. „genau zweimal Seite ‚A‘ bei drei Münzwürfen“) oder Ergebnisberei-

komplexen Streuverhaltens stellen, sind sie u.a. für die Absicherung von Studienergebnissen³⁰², i.R. der industriellen Qualitätssicherung³⁰³, in der beweisfähigen Stichprobenprüfung³⁰⁴ und der automatischen Prozesssteuerung³⁰⁵ anerkannt. Diese Übereinstimmungsprüfung von musterbasierten Häufigkeitsanalysen bedeutet somit die beste dem Menschen bekannte Form der Beurteilung, die zudem die untersuchte Datenmenge im Ergebnis berücksichtigt³⁰⁶.

Sobald musterbasierte Häufigkeitsanalysen in der SRP über die Aufdeckung von Prüffeldern hinaus zur indiziellen Schlüssigkeitsbeurteilung eingesetzt werden, erfolgt grundsätzlich eine Übereinstimmungsbewertung mit Hilfe eines geeigneten Wahrscheinlichkeitstests.

Um die Gefahr ungerechtfertigter Beanstandungen auf Grundlage manipulationsorientierter Mustervergleiche wie Ziffernanalysen möglichst gering zu halten, beschränken sich die Anwendungen i.R. der **SRP** auf **zwei praxistypische Manipulationsarten**:

-  freies Erfinden von Besteuerungsdaten, wodurch Ziffernvorlieben bzw. -abneigungen so deutlich entstanden sein können, dass normale Streutendenzen sie nicht mehr erklären, sowie
-  Spitzeneingriffe als verstärkte Verkürzungen in (betriebsbezogen) hohen Tageskasseneinnahmen, was den manipulationspsychologischen Erfahrungen zu kriminellem Verhalten entspricht (s. [Pkt. 4.3.4.](#)).

Darüber hinaus leitet die SRP mit ihrem nachfolgend abgebildeten „**Prüfungsschema zur Vermeidung von Fehldeutungen**“³⁰⁷ AnwenderInnen durch die manipulationsbezogenen Musteranalysen von Tageskasseneinnahmen, um einerseits vor Anwendungs- sowie Interpretationsfehler zu bewahren und andererseits betriebliche Faktoren als Ursachen für aufgefallene Musterstörungen soweit möglich abzuklären.

che (z.B. „mindestens zweimal Seite ‚A‘ bei drei Münzwürfen“) auf Basis der Wahrscheinlichkeit des einzelnen Vorgangs (für eine ideal geprägte Münze „50:50“) problemlos abgelesen und berechnet werden.

³⁰² Vgl. z.B.: *Bortz/Schuster*, „Statistik für Human- und Sozialwissenschaftler“, Springer-Verlag, 2010, 11.

³⁰³ DIN ISO 2859-1:2014-08: Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung), nachzulesen unter: <https://www.beuth.de/de/norm/din-iso-2859-1/191697831>.

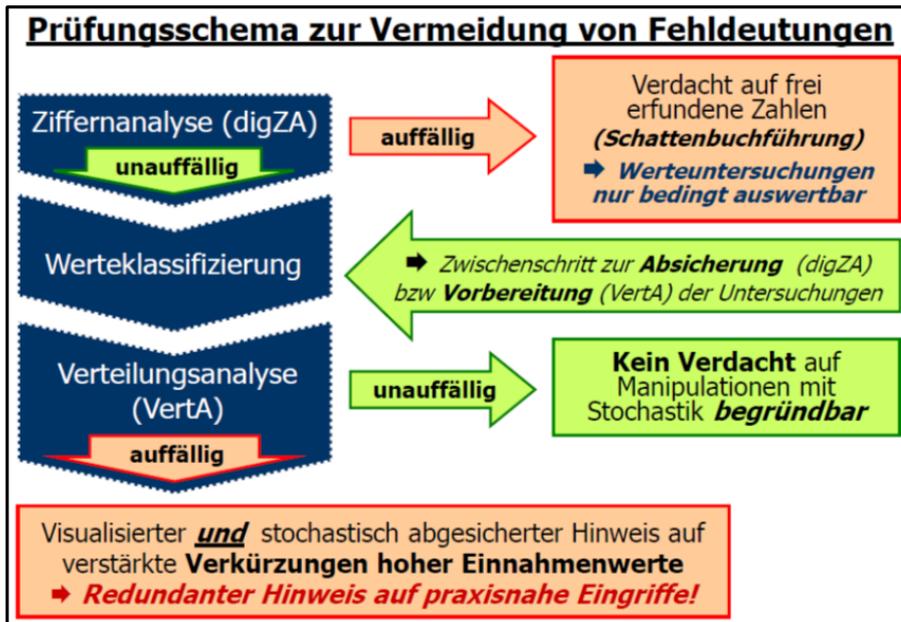
³⁰⁴ § 241 Abs. 1 HGB: „Bei der Aufstellung des Inventars darf der Bestand der Vermögensgegenstände nach Art, Menge und Wert auch mit Hilfe anerkannter mathematisch-statistischer Methoden auf Grund von Stichproben ermittelt werden“ sowie BGH v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 20 ff.

³⁰⁵ Vgl. z.B.: *Richter*, „Prinzipien der Künstlichen Intelligenz“. Teubner-Verlag, 1992.

³⁰⁶ Regelmäßig angemerkte Kritik bei manipulationsorientierten Mustervergleichen wie bspw. der Einerstellenprüfung auf die Gleichverteilung, ob genügend Daten für eine aussagekräftige Beurteilung vorlägen, ist insofern nur angebracht, solange kein Wahrscheinlichkeitstest zur Übereinstimmungsprüfung herangezogen wird. Für einzelne Wahrscheinlichkeitstests kann es Empfehlungen zu Mindestdatenmengen geben wie beim χ^2 -Test (Chi-Quadrat-Test), die jedoch i.d.R. sehr niedrig ausfallen (s. [Pkt. 4.3.1.3.](#)).

³⁰⁷ Das „Prüfungsschema zur Vermeidung von Fehldeutungen“ ist Bestandteil der SRP-Vorlage Stochastik und ergänzt deren Kurzanleitung.

Dazu beginnt die Untersuchung von Tageskasseneinnahmen auf Manipulationsspuren in der SRP grundsätzlich mit den Ziffernanalysen „Einerstelle auf die Gleichverteilung“ und „Zweitiffer auf die Benford-Verteilung“. Hiermit sollen mögliche Ziffernvorlieben bzw. -abneigungen erkannt oder weitgehend ausgeschlossen werden.



Sofern Ziffernauffälligkeiten entdeckt worden sind, werden denkbare betriebliche Ursachen (auch) mit Hilfe der speziellen Funktionalitäten der SRP-Vorlage Stochastik hinterfragt, indem bspw. ein Dezimalsprung in den Daten bei bestimmten Zweitiffernhäufungen mittels einer besonderen Schichtung zügig und einfach als Ursache abgeklärt werden kann (s. das entsprechende Beispiel unter [Pkt. 4.3.2.3.3.](#)). Bevor die kompletten Zahlen der Tageskasseneinnahmen fortfahrend auf stärkere Spitzeneingriffe (in die betriebsbezogenen hohen Einnahmewerte) untersucht werden, überprüft ein Vorbereitungsschritt, ob die Daten einer einzelnen Verteilung folgen. Hierfür stellt die SRP-Vorlage Stochastik zielführende Funktionen zur Verfügung, um bei fehlender Einheitlichkeit einen betrieblichen Grund in Nachfragebesonderheiten der Saison, der Wochentage etc. schnell aufzufinden (s. [Pkt. 4.3.3.3.](#)). Durch den (weitgehenden) Ausschluss von frei erfundenen oder inhomogenen³⁰⁸ Daten kann abschließend eine seriöse Prüfung auf die – kriminalpsychologisch typischen – verstärkten Einnahmenverkürzungen in (betriebsbezogen) hoher Tageskasseneinnahmen vorgenommen werden (s. [Pkt. 4.3.4.](#)).

Durch die weitreichende Vorprogrammierung der SRP-Vorlage Stochastik und das begleitende Prüfungsschema können fehlerleitende Missverständnisse zu manipulationsorientierten Musteranalysen, wie sie auch in der Literatur verbreitet sind³⁰⁹, korrigiert werden. Im Hin-

³⁰⁸ Wenn die Tageskasseneinnahmen z.B. in der Saison oder an Wochenendtagen ein extrem abweichendes Niveau aufweisen, ergeben sich in der grafischen Häufigkeitsauswertung mehrere Häufungsspitzen, so dass eine uneinheitliche (inhomogene) Verteilung entsteht.

³⁰⁹ Vgl. z.B.: *Bisle*, „Chi-Quadrat-Test‘ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PSTR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-

blick auf die allgemeine Akzeptanz von Musteranalysen zur Aufdeckung von Manipulationen und von Wahrscheinlichkeitstests für die Übereinstimmungsprüfung [s. die Abschnitte „musterbasierte Prüfungsansätze zur Aufdeckung von Manipulationen“ und „Wahrscheinlichkeitslehre (Stochastik)“ zu diesem Pkt. mit entsprechenden Verweisen] kann die derzeitige Zurückhaltung in der Finanzrechtsprechung³¹⁰ nur darauf zurückzuführen sein, dass Fehler in der Anwendung befürchtet werden³¹¹. Denn in der Literatur wird zu Recht auf den qualifizierten wissenschaftlichen Hintergrund der Verfahren hingewiesen: „Erwähnenswert ist in diesem Zusammenhang eine Untersuchung der Universität Münster (Institut für Wirtschaftsinformatik) über die Aussagekraft des Chi-Quadrat-Tests. [...] Diese Untersuchung dokumentiert deutlich, dass diese statistische-mathematische Prüfungsmethode auf wissenschaftlich gesicherten Grundlagen beruht.“³¹²

4.3.1. Ablauf von Häufigkeitsanalysen (Chi-Quadrat-Test)

Musterbasierte Häufigkeitsanalysen folgen in ihrem Ablauf allgemein diesem **Schema**:



4.3.1.1. Beobachtungen i.R. von Häufigkeitsanalysen

Für Beobachtungen der Eigenschaften von (Besteuerungs-)Daten kommen alle **skalierbaren Informationen** in Betracht. Damit sind Datenmerkmale³¹³ gemeint, welche bereits

Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Erkenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.

³¹⁰ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. Trossen; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. Zimmermann [Fahrtbuch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. Kühnen; Sosna, StBp. 2000, 41 [Teil I], 68 [Teil II]; Meyer, DStR 2005, 2114; Watrin/Stuffert, DB 2006, 1748; Huber, StBp. 2009, 121 [124 ff.]“.

³¹¹ Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

³¹² Madauß, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 15.

³¹³ Die Datenmerkmale befinden sich in einer Datenbank zu jedem Datensatz vereinheitlicht auf die Spalten verteilt.

diskret³¹⁴ differenzierbar wie der USt-Satz, die Zahlungsweise, die Kostenart etc. sind oder als stetige³¹⁵ Wertinhalte in Klassen unterteilt werden können. Nach dieser Eigenschaft werden die musterbasiert zu untersuchenden Daten ausgezählt. Das Ergebnis kann tabellarisch oder in einem Diagramm dargestellt werden: die unterschiedlichen Merkmalsausprägungen³¹⁶ auf der X-Achse und die jeweiligen Häufigkeiten auf der Y-Achse (*zur Überlegenheit von verbildlichten Analyseergebnissen s. den Abschnitt „Visualisierung von Besteuerungsdaten“ unter [Pkt. 2.](#)*).

Im Hinblick auf den Wertinhalt von bspw. Tageskasseneinnahmen können als *eine* Eigenschaft die Ziffern nach ihren Häufigkeiten ausgezählt werden. Hierbei ist die Betrachtungsrichtung zu unterscheiden: von vorne mit den Zahlenpositionen „Erstziffer“, „Zweitziffer“, „Drittziffer“ usw. oder mit vergleichbarer Größenordnung als „ Einerstelle“, „Zehnerstelle“, „Hunderterstelle“ etc. Die erstgenannte Zahlenbeschreibung, auf der die Verteilungserfahrungen des Benford-Gesetzes (s. [Pkt. 4.3.2.3.1.](#)) basieren, kann dazu führen, dass Zahlenpositionen mit abweichender Wertigkeit innerhalb eines Datenbestands gemeinsam ausgewertet werden, indem z.B. die Zweitziffer der obigen Tageskasseneinnahme „€ 1.276,20“ die *Hunderterstelle* meint, während an einem absatzschwächeren Tag mit „€ 699,30“ die *Zehnerposition* die zweite Stelle von vorne bildet. Analysebesonderheiten, die sich aus der Überschneidung beider Beschreibungsperspektiven ergeben, werden von der SRP-Vorlage Stochastik automatisch beachtet (s. [Pkt. 4.3.2.2.1.](#)).



Weiter können die Datensätze in Größenklassen eingeteilt werden. Um gleichmäßige Muster erkennen zu können, werden dabei i.d.R. konstante³¹⁷ Klassenabstände gewählt wie in der obigen Abb. mit dem € 100-Intervall „> 1.200 – 1.300“. Die SRP-Vorlage Stochastik führt derartige Datenklassifizierungen (Schichtungen) automatisch auf Basis der praxiserprobten Wurzelmethode³¹⁸ durch, bietet den AnwenderInnen auf den Tabellenblättern

³¹⁴ Unter diskreten Verteilungen werden in der Mathematik solche verstanden, die keine Zwischenwerte einnehmen können. So kann es bei Münzwürfen aus Sicht einer bestimmten Münzseite nur zu ganzzahligen Erfolgen (bzw. Misserfolgen) kommen, nicht aber zu 3,437 Treffern. Das Gegenteil mit kontinuierlichen Ergebnismöglichkeiten wird „stetig“ genannt.

³¹⁵ Vgl. die Erklärung zu „diskret“ unter der vorherigen Fußnote.

³¹⁶ Zu jedem differenzierbaren Datenmerkmal – in der Datenanalytik auch Dimension genannt – gibt es die originären oder festgelegten (in Klassen unterteilten) Ausprägungen wie bspw. zum USt-Satz „0 %“, „7 %“, „19 %“ etc.

³¹⁷ Sofern ein logarithmisches Muster wie die Logarithmische Normalverteilung erwartet wird, ist die Wahl von Größenklassen mit logarithmisch gleichen Abständen (z.B. „10²“ – „10³“ – „10⁴“ – „10⁵“ etc. entsprechend „100“ – „1.000“ – „10.000“ – „100.000“ usw.) oft zielführender, um die Gleichmäßigkeit der beobachteten Verteilung visuell zu beurteilen.

³¹⁸ Die Wurzelmethode berechnet die Klassenanzahl zur Klassifizierung (Schichtung) nach folgender Formel: $\sqrt{\text{Anzahl Datensätze}}$. Das Ergebnis wird in der SRP-Vorlage Stochastik zur besseren Übersichtlichkeit auf max. 30 Intervalle begrenzt.

„Schichtung“, „Schichtungsvertiefung“ und „Verteilungsanalyse“ jedoch alle Optionen zur freien Anpassung der Schichtobergrenzen – und damit Intervalle – sowie zur Umstellung von linear auf logarithmisch gleichmäßige Klassen an.

Darüber hinaus liefert eine vorhandene Chronologie in zu untersuchenden Daten weitere Beobachtungsoptionen mit dem Wiederholungsverhalten (Iteration) an: Ausgewertet wird bspw. wie viele Datensätze es dauert, bis die beobachtete Einerstelle „6“ erneut vorkommt. Dieser Ansatz ist insbesondere für moderne (digitale) Einnahmenverkürzungen [s. den Abschnitt „raffinierte (digitale) Manipulationswerkzeuge“ unter [Pkt. 2.1.](#)] von Bedeutung, weil die zeitliche Ziffernabfolge, welche von Zufallsgeneratoren erzeugt wird, nicht derjenigen entspricht, die natürlich entsteht.

4.3.1.2. (Punkt-)Erwartungen i.R. von Häufigkeitsanalysen

Zu jeder Schlussfolgerung über eine Beobachtung von Datenmerkmalen braucht es ein **Vergleichsmaß**. Diese Erwartung stellt als Punktideal³¹⁹ die Grundlage für eine Beurteilung des beobachteten Häufigkeitsmusters im betriebswirtschaftlichen, rechtlichen oder einem anderen Kontext.

Erwartungsmuster können von Erfahrungen aus dem geprüften Unternehmen stammen (interner Strukturvergleich) wie z.B. die Häufigkeitsverteilung von Kostenpositionen aus vorherigen Prüfungszeiträumen oder im Vergleich mehrerer vergleichbarer Betriebe eines Konzerns. Im Fall deutlich erkennbarer Unterschiede im Verteilungsmuster betriebs- oder konzernbezogener Daten werden dadurch wesentlich veränderte Verhältnisse angezeigt, die auf umgestellte Handlungsabläufe, Umstrukturierungen oder menschliche Eingriffe hinweisen und somit einzelfallspezifische Prüffelder offenlegen.

Darüber hinaus sind allgemeingültige Verteilungsmuster, die in (Wirtschafts-)Daten regelmäßig vorgefunden wurden, bekannt und können für Vergleiche herangezogen werden (externer Strukturvergleich). Das bekannteste Häufigkeitsmuster, welches wegen seiner weiten Verbreitung auch dann angewendet wird, wenn kein anderes bekannt ist, stellt die Normalverteilung, wie sie z.B. die Verhältnisse zum Intelligenzquotienten in (Teilen) der Bevölkerung beschreibt (s. die Abb. zum Abschnitt „gleichmäßige Muster der Häufigkeiten“ unter [Pkt. 4.3.](#)). Sie dient auch als Grundlagenverteilung für diverse Wahrscheinlichkeitstests, weil sie als typisches natürliches Streuverhalten gilt (s.a. [Pkt. 4.3.1.3.](#)). Mit ihrer ebenfalls symmetrischen Glockenform ist die Binomialverteilung³²⁰ ein der Normalverteilung ähnliches Muster und stellt sich bei (dichotomen) ja-nein-Sachverhalten wie bspw. dem Verhältnis

³¹⁹ Die ideale Punkterwartung kennt zunächst keine Toleranz wie z.B. ein exaktes Rechenergebnis aus einer Nachkalkulation. Erst i.R. der Übereinstimmungsprüfung werden Unsicherheiten der Berechnung oder aus unvollständigen Informationen als tolerierbarer Ergebnisrahmen einbezogen.

³²⁰ Die Binomialverteilung kennt als diskrete Form im Gegensatz zur Normalverteilung keine Zwischenwerte und wird somit als Säulendiagramm wiedergegeben.

„Verzehr an Ort und Stelle zu Außerhausverkauf“ um das Standardverhältnis³²¹ herum ein. Daneben gibt die Gleichverteilung mit identischen Anteilen für alle möglichen Merkmalsausprägungen das Abbild des absoluten, gleichgewichteten Zufalls und ist somit als weitere generelle Erwartung geeignet (s. die beiden Zifferanalyse-Grafiken im Abschnitt „Übereinstimmung mit dem Idealmuster von der Datenmenge abhängig“ unter [Pkt. 4.3.](#)). Nach empirischen Erkenntnissen von Newcomb und Benford weisen besonders die führenden Zahlenpositionen in Wirtschaftsdaten weitere gleichmäßige Häufigkeitsmuster der (möglichen) Ziffern auf (s. [Pkt. 4.3.2.3.1.](#)).³²² In vielen Wissenschaftsbereichen wird zudem ein homogenes rechtsschiefes Muster für die nach Größenklassen ausgezählten Daten wiedergefunden, wenn diese in ihrer Entfaltung linksseitig begrenzt³²³ sind: die Logarithmische Normalverteilung³²⁴. Unter logarithmischer³²⁵ Skalierung der X-Achse wird aus jeder Logarithmischen Normalverteilung eine Normalverteilung (s. [Pkt. 4.3.4.1.](#)).

Den genannten allgemeinen Verteilungsmustern ist gemein, dass sie mit Hilfe weniger Informationen³²⁶ vollständig beschrieben und deshalb problemlos auf die speziellen Häufigkeitsverhältnisse der untersuchten (Besteuerungs-)Daten angepasst werden können. In der SRP-Vorlage Stochastik werden die erwarteten Häufigkeitsmuster automatisch und nachvollziehbar aus den eingepflegten Daten auf Basis der jeweiligen Verteilungsfunktion gebildet. Für die manipulationsorientierte Analyse der Wertzusammenhänge in Tageskasseneinnahmen auf den Tabellenblättern „Verteilungsanalyse“ und „Verteilungsdiagramm“ – im Anschluss an die Vorbereitungsmaßnahmen gem. SRP-Prüfungsschema (s. den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)) – ist eine gezielte Anpassung der Punkterwartung „Logarithmische Normalverteilung“ möglich. Auf diese Weise

³²¹ Im Vergleich mehrerer Filialen mit Erlöstrennung in „Verzehr an Ort und Stelle“ und „Außerhausverkauf“ hat jede ein typisches Nachfrageverhältnis. Werden bspw. die einzelnen Tage nach den Relationen ausgezählt, ergibt sich für jede Filiale (ohne wesentliche strukturelle Veränderungen) eine Binomialverteilung um den jeweiligen typischen Verhältniswert herum.

³²² Benford, „The Law of Anomalous Numbers“, Proceedings of the American Philosophical Society, 1938, 551.

³²³ Als praxisnahes Beispiel seien hier die Kraftstoffverbräuche von Pkw-Typen genannt. Unterhalb eines minimalen Verbrauchs sind technisch keine Autos verfügbar, was die linksseitige Begrenzung bedeutet. Mit auf der X-Achse ansteigendem Kraftstoffverbrauch nehmen die Pkw-Typen bis zu einem (für den Beobachtungszeitpunkt) typischen Normverbrauch zu. In den darüber hinausgehenden Verbrauchsklassen lassen die Häufigkeiten langsamer (als linksseitig des Normverbrauchs) nach und reichen erheblich weiter nach rechts bis zum Verbrauchmaximum.

³²⁴ Vgl. z.B.: *Limpert/Stahel/Abbt*, „Lognormal distributions across the sciences: keys and clues“, BioScience Band 51, Nr. 5, 2001, 341, nachzulesen unter: https://de.wikipedia.org/wiki/Logarithmische_Normalverteilung.

³²⁵ Eine logarithmische Skalierung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie bspw. mit „10⁰ – 10¹ – 10² – 10³“ (entsprechend „1 – 10 – 100 – 1.000“).

³²⁶ Die Berechnung der Häufigkeitsanteile für die Normal- und die Logarithmische Normalverteilung erfordert den Mittelwert und die Standardabweichung (vorzeichenneutrale durchschnittliche Abweichung jedes einzelnen Datensatzes vom Mittelwert). In der Binomialverteilung lassen sich die Verhältnisse aus der Trefferanzahl, Trefferwahrscheinlichkeit und Datenanzahl errechnen. Für die anteiligen Häufigkeiten nach den Benford-Verteilungen wird die betreffende Ziffer benötigt.

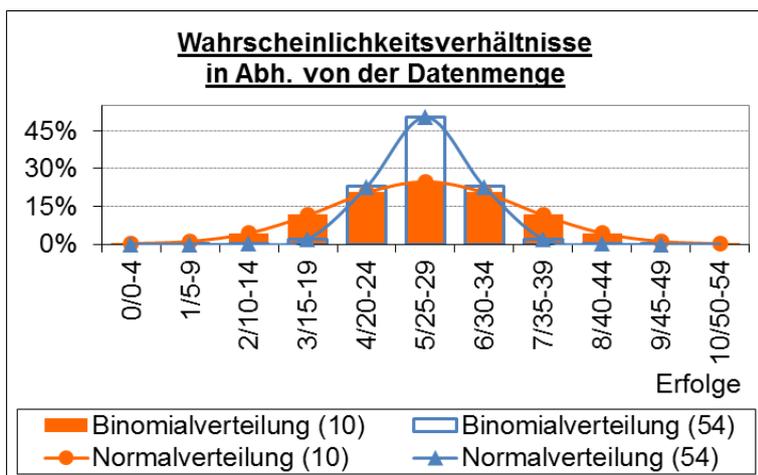
kann der Verdacht eines – manipulationspsychologisch typischen – stärkeren Spitzeneingriffs einzelfallbezogen untersucht werden (s. [Pkt. 4.3.4.3](#)).

4.3.1.3. Übereinstimmungsprüfung i.R. von Häufigkeitsanalysen

Wie bei jedem prüfenden Soll-Ist-Vergleich schließt sich an eine musterbasierte Häufigkeitsbeobachtung zu den (Besteuerungs-)Daten und der Zuordnung einer an den Einzelfall angepassten Punkterwartung die Übereinstimmungsprüfung an.

Da eine genaue Übereinstimmung mit der Mustererwartung nur in seltenen Fällen eintreten wird, umfasst die Übereinstimmungsprüfung auch die **Festlegung eines Unbedenklichkeitsbereichs**. Vergleichbar wird bspw. beim externen Richtsatzvergleich der beobachtete Aufschlagsatz nicht am mittleren Richtsatzwert sondern an der Richtsatzspanne gemessen und bei einer Nachkalkulation ein Ergebnisrahmen durch Hinzu- und Abrechnung eines bestimmten Betrags bzw. Prozentanteils definiert. Dieser besonders bei komplexen Gesamtbildprüfungen übliche Toleranzrahmen soll die Informationslücken und Berechnungsunsicherheiten berücksichtigen. I.R. von musterbasierten Häufigkeitsanalysen, welche für indizielle Aussagen gem. § 158 AO gedacht sind, werden hierfür i.d.R. die Mittel des mathematischen Spezialgebiets für „Aussagen unter Unsicherheit“ verwendet: die **Wahrscheinlichkeitstests**.

Im Gegensatz zu anderen Soll-Ist-Vergleichen der Betriebsprüfung ist das schon deshalb bei Häufigkeitsanalysen zu empfehlen, weil diese Verfahren die Abhängigkeit der Mustertreue von der Anzahl untersuchter Daten (s. *das Beispiel im Abschnitt „Übereinstimmung mit dem Idealmuster von der Datenmenge abhängig“* unter [Pkt. 4.3](#)) beachten. So lässt sich



z.B. an den nebenstehenden Binomialverteilungen problemlos ablesen, dass bei **zehn Würfeln mit einer ideal geprägten³²⁷ Münze** ein sehr viel breiteres Ergebnisspektrum um die Punkterwartung „5-mal jede Seite“ resultiert als bei **deutlich mehr Versuchen (54)**. Auch die Normalverteilung als Grundlage

diverser Wahrscheinlichkeitstests beinhaltet die Eigenschaft, dass die Streuungsintensität von der Datenanzahl abhängt. In der Konsequenz berücksichtigt eine musterbasierte Häufigkeitsanalyse, welche die Übereinstimmungsprüfung mit einem Wahrscheinlichkeitstest vor-

³²⁷ Damit ist die Ergebniswahrscheinlichkeit für die Seiten „A“ und „B“ für jeden einzelnen Wurf „50:50“.

nimmt, die untersuchte Datenmenge: Werden bspw. wenige Tageskasseneinnahmen aus einer abgekürzten Außenprüfung auf die Übereinstimmung ihrer Einerstelle mit der Gleichverteilung untersucht, erwartet der χ^2 -Test (**Chi-Quadrat-Test**) eine wesentlich schwächere Deckungsgleichheit als bei den umfangreicheren Daten einer normalen Betriebsprüfung. Insofern sind die in der Öffentlichkeit mitunter wahrzunehmenden Forderungen nach hohen Mindestdatenmengen wie z.B. „ ≥ 500 “ unberechtigt. Vielmehr gibt es bspw. zum χ^2 -Test (Chi-Quadrat-Test) eine vorherrschende Empfehlung, wie viele Daten für eine seriöse Auswertung mindestens vorliegen sollten³²⁸, und das sind bei den typischen Anwendungsformen deutlich weniger: So bedeutet ein Mindestdatenumfang von „ ≥ 5 erwarteten Elementen pro Klasse“ für die Gleichverteilung zehn möglicher Ziffern, dass mindestens 50 Datensätze zur Untersuchung vorliegen sollten. Bei einem Mustervergleich mit den Benford-Verteilungen richtet sich die Mindestdatenmenge nach dem kleinsten Erwartungsanteil, was für die Zweit-ziffer mit ca. 8,5 % für die „9“ eine Empfehlung von minimal 59 Datensätzen ergibt.

Auf Basis der Verteilungen, die den Wahrscheinlichkeitstests zu Grunde liegen, werden die Ergebnisse ihrer (stochastischen) Übereinstimmungsprüfung als prozentuale Übereinstimmungs- bzw. Nichtübereinstimmungswahrscheinlichkeit ausgegeben. Die Summe aus beiden ergibt grundsätzlich 1 (100 %). Im Gegensatz zu bspw. „Kalkulationen“ oder „Zeitreihenanalysen“ resultiert somit aus der stochastischen Übereinstimmungsbeurteilung eines beobachteten Häufigkeitsmusters mit der (auf den Einzelfall angepassten) Punkterwartung ein konkretes Ergebnis, welches transparent und jederzeit reproduzierbar ist. Demzufolge wird diese Form der Beurteilung unter Unsicherheit grundsätzlich anerkannt wie z.B. in der Absicherung von Studienergebnissen³²⁹, der industriellen Qualitätssicherung³³⁰, der beweisfähigen Stichprobenprüfung³³¹, der automatischen Prozesssteuerung³³² usw.

In der Praxis haben sich einzelne Wahrscheinlichkeitstests für bestimmte Fragestellungen durchgesetzt, so dass nur sie in Kalkulationsprogrammen wie MS Excel®, OpenOffice® etc. enthalten sind. Sollen zwei Verteilungen mit mindestens drei Klassen wie bei der Ziffernanalyse mit neun bzw. zehn möglichen Ziffern auf Übereinstimmung geprüft werden, ist der Standardtest der χ^2 -Test (Chi-Quadrat-Test). Gibt es nur zwei Ergebnisklassen wie z.B.

³²⁸ Diller/Schmid/Späth/Kühne, „Zifferntests in der Betriebsprüfung – Chancen und Risiken“, DStR 6/2015, 311, 314: „Als Mindestumfang [für den Chi²-Test] werden in der Literatur unterschiedliche Werte angegeben, wobei die am häufigsten vorgeschlagene Regel besagt, dass jede Klasse im Erwartungswert mit mindestens 5 Elementen besetzt sein sollte“ mit Verweis auf Rüger, „Test- und Schätztheorie“, 2002, 204.

³²⁹ Vgl. z.B.: Bortz/Schuster, „Statistik für Human- und Sozialwissenschaftler“, Springer-Verlag, 2010, 11.

³³⁰ DIN ISO 2859-1:2014-08: Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung), nachzulesen unter: <https://www.beuth.de/de/norm/din-iso-2859-1/191697831>.

³³¹ § 241 Abs. 1 HGB: „Bei der Aufstellung des Inventars darf der Bestand der Vermögensgegenstände nach Art, Menge und Wert auch mit Hilfe anerkannter mathematisch-statistischer Methoden auf Grund von Stichproben ermittelt werden“ sowie BGH v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 20 ff.

³³² Vgl. z.B.: Richter, „Prinzipien der Künstlichen Intelligenz“. Teubner-Verlag, 1992.

beim Münzwurf oder werden die möglichen Ausgänge in zwei Klassen eingeteilt³³³ ist für diese (dichotome) Fragestellung die **Binomialverteilung** einschlägig. Aufgrund der allgemeinen Verbreitung und z.T. kostenlosen Zugänglichkeit von Tabellenkalkulationsprogrammen können die Ergebnisse von musterbasierten Häufigkeitsanalysen von jedem „sachverständigen Dritten“³³⁴ problemlos reproduziert werden. Dabei hat sich die Anwendung mit Komplettfunktionen wie bspw. „CHIQU.TEST“ in MS Excel®, welche nur noch die beiden Datenreihen im Kontextmenü zugewiesen werden müssen, gegenüber den früher üblichen Ergebnistabellen erheblich erleichtert.

I.R. der SRP werden alle manipulationsorientierten Musteranalysen aufgrund des indiziellen Charakters mit einem für die jeweilige Fragestellung gängigen Wahrscheinlichkeitstest auf Übereinstimmung geprüft. Zur besseren Transparenz werden beim χ^2 -Test (Chi-Quadrat-Test) auch die Zwischenergebnisse „ χ^2 -Klassenwert“ und „ χ^2 -Gesamtwert“ ausgegeben. Die Anwendungsempfehlung zur Mindestmenge an die zu untersuchenden Daten „ ≥ 5 Erwartungen pro Klasse“ wird automatisch und durchgehend in der SRP-Vorlage Stochastik umgesetzt. Sobald – insbesondere infolge von Dateneingrenzungen auf dem Tabellenblatt „Daten“ (s. den Abschnitt „zweistufige Eingrenzung der untersuchten Daten“ unter [Pkt. 4.](#)) – die empfohlene Mindestdatenmenge unterschritten wird, werden entweder die Testergebnisse ausgegraut (s. [Pkt. 4.3.2.1.](#)) oder die betreffenden Klassen in der Auswertung nicht berücksichtigt³³⁵ (s. [Pkt. 4.3.4.2.](#)). Als Hinweis auf ein auffälliges Häufigkeitsmuster erfolgt auf den Auswertungsblättern „digZA“ und „Verteilungsanalyse“ der SRP-Vorlage Stochastik eine Hervorhebung der Testergebnisse in den zwei Stufen „ $\geq 99\%$ “ durch rote Hintergrundfarbe und „ $\geq 99,95\%$ “ zusätzlich durch rote Schrift. Während die erste Auffälligkeitsschwelle – nach erfolgloser Abklärung besonderer einzelbetrieblicher Ursachen – für das Vorliegen mehrerer Indizien als Indiziengesamtheit³³⁶ (s. [Pkt. 4.5.](#)) gedacht ist, gilt die zweite strengere für Einzelindizien.

³³³ Bei Ziffernanalysen wird aus der (visualisierten) Gegenüberstellung erkennbar, ob es Auffälligkeiten einzelner Ziffern geben könnte. Soll die Untersuchung im Hinblick auf eine einzelne zu häufig oder selten vorkommende Ziffer vertieft werden, resultieren daraus die möglichen Ergebnisse „Ziffer x“ und „alle anderen Ziffern“, also eine (dichotome) ja-nein-Verteilung mit nur noch zwei Klassen. Die Idealwahrscheinlichkeiten richten sich dann nach der jeweilig zutreffenden Punkterwartung mit z.B. „10 %“ zu „90 %“ bei der Gleichverteilung.

³³⁴ § 145 Abs. 1 AO.

³³⁵ Kurvenverteilungen, die sich (asymptotisch) der X-Achse annähern, wie die Normal- und die Logarithmische Normalverteilung fallen zwangsweise zu den Seiten hin irgendwann unter die Klassengrenze „ ≥ 5 “. Deshalb wird der χ^2 -Test (Chi-Quadrat-Test) als Übereinstimmungsprüfung auf die Logarithmische Normalverteilung i.R. der Verteilungsanalyse auf die Klassen mit einer Erwartung „ ≥ 5 “ beschränkt (s. [Pkt. 4.3.4.2.](#)).

³³⁶ BGH v. 27.4.2010 – 1 StR 454/09, NSTZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweismwürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die

4.3.1.4. Schlussfolgerungen i.R. von Häufigkeitsanalysen

Die aus einem Übereinstimmungsergebnis wie z.B. 99 % Nichtübereinstimmung des Einerstellenmusters aus Tageskasseneinnahmen mit der Gleichverteilung abzuleitende korrekte Schlussfolgerung ist, dass diese Besteuerungsdaten mit „sehr hoher Wahrscheinlichkeit“ (s. [Pkt. 4.6.](#)) nicht unbeeinflusst von menschlichen Eingriffen entstanden sind.

Um diese Erkenntnis als beweiskräftiges **Indiz** für frei erfundene Bareinnahmen heranziehen zu können, muss zunächst noch (weitestgehend) ausgeschlossen werden, dass die beiden anderen Formen menschlicher Beeinflussung „Fehler bei der Erfassung“ und „Gestaltungen“ ursächlich sind. Die Gefahr fehlerhafter Besteuerungsdaten wie z.B. tlw. für mehrere Tage zusammengefasste Kassenerlöse kann mit Hilfe von Datenbankenfunktionen wie Filtern, Sortieren etc. zügig abgeklärt und das Ergebnis unter Angabe der verwendeten Daten nachvollziehbar dokumentiert werden. Der Einfluss möglicher Gestaltungen lässt sich bereits deshalb i.d.R. unproblematisch hinterfragen, weil Tageskasseneinnahmen lediglich von den zwei Faktoren „Nachfragemenge(n)“ und „Einzelpreis(e)“ abhängig sind: So können ausschließlich auf „0“ und „5“ Euro endende Einzelpreise an der Einerstelle in Tagessummen auch nur dieselben Ziffern hervorbringen. Als weiteres Beispiel werden Sondertagesveranstaltungen häufig mit einem Pauschalpreis abgegolten, so dass der Nachfragezufall keine freie Entstehung der Ziffernhäufigkeiten bewirken kann. Insgesamt ist für die i.R. der SRP durchgeführten manipulationsorientierten Musteranalysen ein (weitest gehender) Ausschluss anderer Gründe als manipulativer für eindeutig abweichende Beobachtungen ohne mathematische Spezialkenntnisse und auf der Basis von (einigermaßen) ordnungsgemäßen Aufzeichnungen möglich.³³⁷

Nach dem SRP-Prüfungsschema (s. *den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)*) werden auffällige Ergebnisse manipulationsorientierter Musteranalysen selbständig auf die Verursachung durch die (wenigen) praxisrelevanten betrieblichen Faktoren hinterfragt. Dabei helfen in schwierigeren Fragen die Funktionalitäten der SRP-Vorlage Stochastik, wenn z.B. abzuklären ist, ob eine typische Zweitiffernhäufung auf einen in den Daten vorliegenden Dezimalsprung zurückgeführt werden kann³³⁸ (s. *das entsprechende Beispiel unter [Pkt. 4.3.2.3.3.](#)*). Sofern auf diese Weise die anderen Formen menschlicher Einflüsse auf die Tageskasseneinnahmen (weitest gehend) ausgeschlossen

entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

³³⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Dabei ist allerdings die sphärenorientierte Mitwirkungspflicht des Stpfl. (s. § 88 Rz. 3 f.; § 90 Rz. 20) zu beachten. Mathematisch eindeutige Unschlüssigkeiten der Buchführungsergebnisse muss er durch betriebsspezifische Besonderheiten schlüssig erklären“.

³³⁸ Die weiteren praxisrelevanten betrieblichen Besonderheiten, die ein auffälliges Häufigkeitsmuster der Einerstelle, Zweitiffer oder Werteverteilung verursacht haben können, werden unter den jeweiligen Abhandlungen der Analysen beschrieben (s. *Pkt. 4.3.2.2.3., Pkt. 4.3.2.3.3. sowie Pkt. 4.3.4.5.*).

werden können, erreicht die auffällige Beobachtung den Status eines Indizes gegen die sachliche Richtigkeit mit jener Beweiskraft, die der Wahrscheinlichkeitstest zur Nichtübereinstimmung ausgegeben hat.

Bisher ist die Akzeptanz manipulationsorientierter Musteranalysen in der Finanzrechtsprechung (noch) sehr zurückhaltend³³⁹. Ursache dafür dürften die Komplexität der Methoden und Tatsachen verzerrende Veröffentlichungen³⁴⁰ sein. Deshalb wurde in der SRP von Beginn an sehr großen Wert auf die Unterscheidung der einzelnen Analyseschritte und deren korrekte Beschreibung gelegt.

4.3.2. Ziffernanalyse

(Digitale) Ziffernanalysen verfolgen das Ziel, das **freie Erfinden von Besteuerungsdaten** über Spuren aufzudecken, die infolgedessen unbewusst hinterlassen wurden: Vorlieben für bzw. Abneigungen gegen bestimmte Ziffern oder Ziffernchaos. Untersuchungen zum Ausdenken von Zahlen haben gezeigt, dass dabei häufiger einzelne Ziffern so stark bevorzugt oder unterdrückt wurden, dass sie außerhalb der zufällig möglichen Häufigkeiten vorkamen.³⁴¹ Weil das natürliche Streuverhalten (um ein typisches Ergebnis herum) dem mathematischen Teilgebiet der Wahrscheinlichkeitslehre (Stochastik) so genau bekannt ist (s. [Pkt. 4.3.1.3.](#)), können zu auffälligen Ziffernkonstellationen konkrete (prozentuale) Aussagen getroffen werden, wie sicher es ist, dass die untersuchten (Besteuerungs-)Daten menschlich beeinflusst worden sein müssen. Der Umkehrschluss, dass bspw. Tageskasseneinnahmen nicht frei erfunden worden sind, lässt sich mit Ziffernanalysen jedoch nicht qualifiziert bestätigen.

³³⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. *Trossen*; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. *Zimmermann* [Fahrtbuch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. *Kühnen*; *Sosna*, StBp. 2000, 41 [Teil I], 68 [Teil II]; *Meyer*, DStR 2005, 2114; *Watrin/Stuffert*, DB 2006, 1748; *Huber*, StBp. 2009, 121 [124 ff.]; krit. gegenüber der zurückhaltenden FG-Rspr.: *Wähnert*, StBp. 2011, 269 [271]; *Wähnert*, StBp. 2012, 241 [244]; *Wähnert*, Stbg. 2014, 20 [22 f.]“

³⁴⁰ Vgl. z.B.: *Bisle*, „Chi-Quadrat-Test‘ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PStR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Erkenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.

³⁴¹ *Hill*, „The Difficulty of faking Data“, *Chance* 12/3/1999, 27, sowie *Hill*, „Random-Number Guessing and the First Digit Phenomenon“, *Psychological Reports* 62/1988, 967.

Das wichtigste **Manipulationsszenario**, welches mit Ziffernanalysen aufgedeckt werden kann, ist die manuelle Aufzeichnung i.d.R. von Tageskasseneinnahmen. Darüber hinaus können z.B. bei den folgenden moderneren (digital unterstützten) Eingriffsformen ebenfalls eindeutige Ziffernvorlieben oder -abneigungen als „Fingerabdrücke“ in die Ziffernverteilungen gelangen:

-  Beim „Zwei-Kassen-Modell“ („Double-Till“) wird zusätzlich zur Geschäftskasse mit den korrekten Erfassungen (für die Personalkontrolle) eine zweite Kasse desselben Typs im Verborgenen mit Phantasiezahlen gespeist. Um den Arbeitsaufwand im erträglichen Rahmen zu halten, werden meistens Tagessummen für die relevanten Berichtspositionen eingegeben und verräterische Kassenberichtsangaben unterdrückt. Somit kann es zum freien Erfinden der steuerlich aufgezeichneten Tageskasseneinnahmen kommen.
-  Die bei älteren Registrierkassenmodellen nach Prüfungserfahrungen verbreitetere Praxis, zunächst per X-Abschlag (ohne Nullstellung) die Bareinnahmen zu kontrollieren, anschließend mittels Manager-Storno zu löschen und einen Wunschbetrag für die Besteuerung einzugeben, kann auffällige Ziffernhäufigkeiten hervorrufen.
-  Selbst einige Varianten systematischer digitaler Manipulation beinhalten die Eingabe eines „gewünschten“ Betrags für die Tageskasseneinnahmen. So sahen vor allem frühere Zapperversionen³⁴² für Warenwirtschaftsprogramme in Apotheken vor, dass die „neue“ Tageskasseneinnahme in ein bestimmtes Feld eingegeben wird. Danach erfolgte die intelligente Umverteilung auf die im Bericht auszugebenden Positionen wie die Aufteilung nach USt-Sätzen, Zahlungswegen etc.

Demgegenüber können Manipulationsweisen, die auf dem Löschen von Positionen oder auf einer prozentualen Reduktion von Werten beruhen, nicht durch Ziffernanalysen aufgedeckt werden, weil die verkürzten Tageskasseneinnahmen weiterhin ein Produkt zufälliger Entstehung sind.

Tipp: Für das grundsätzliche Verständnis manipulationsorientierter Musteruntersuchungen ist das Verständnis der Ziffernanalysen hilfreich, weil daraus die Übertragung der einzelnen Annahmen und Schritte des allgemeinen Ablaufs (s. [Pkt. 4.3.1.](#)) auf komplexere Ansätze wie insbesondere die wertbezogene Verteilungsanalyse (s. [Pkt. 4.3.4.](#)) leichter fällt.

4.3.2.1. Anwendungshinweise zu SRP-Ziffernanalysen

Folgende Hinweise zur Durchführung von Ziffernanalysen mit der SRP-Vorlage Stochastik sind wichtig:

³⁴² Unter „Zappern“ werden externe Manipulationstools für elektronische Kassen(-systeme) verstanden, die als Programme z.B. auf einem USB-Stick liegen. Nach Ankopplung an das Aufzeichnungssystem gewähren sie den Lösch- und Änderungszugriff auf die gespeicherten Daten.

- ➔ Sobald die zu untersuchenden Daten in die beiden Spalten „Datum“ und „alle Werte“ auf dem Tabellenblatt „Daten“ hineinkopiert wurden, werden die Einerstelle³⁴³ und Zweitzi-fer auf dem Tabellenblatt „digZA“ automatisch nach ihren Mustern ausgewertet und den jeweiligen (Punkt-)Erwartungen (s. [Pkt. 4.3.1.2.](#)) „Gleich-“ bzw. „Benford-Verteilung“ ge-gegenübergestellt. Der **redundante Ansatz** mit zwei untersuchten Zahlenstellen soll ei-ne bessere Aufdeckungswahrscheinlichkeit für Ziffernauffälligkeiten erreichen, weil viele Manipulationsspuren im weiten Streubereich des Zufalls untergehen.
- ➔ Beide (kompletten) Ziffernanalysen werden mit einem χ^2 -Test (**Chi-Quadrat-Test**) auf Übereinstimmung geprüft (s. [Pkt. 4.3.1.3.](#)). Wird die empfohlene Mindestdatenmenge „ ≥ 5 Elemente pro Erwartungsklasse“ unterschritten, verfärben sich die betroffenen Fel-der der Erwartungsspalte rot und wird das χ^2 -Testergebnis mit hellgrauer Schrift unter-drückt³⁴⁴. Mathematisch eindeutige (signifikante) Abweichungen vom jeweiligen Erwar-tungsmuster werden ab 99 % Gegenwahrscheinlichkeit mit roter Feldfarbe und ab 99,95 % Gegenwahrscheinlichkeit zusätzlich mit roter Schrift hervorgehoben. Außerdem erfolgt folgender (per Anhakfeld ausblendbare) Hinweis: „Betriebliche Ursachen sollten abgeprüft werden.“
- ➔ Fallen in den Gegenüberstellungen der beobachteten und erwarteten Ziffernhäufigkeiten einzelne Ziffern als besonders häufig oder selten auf, werden sie in der betreffenden Vergleichstabelle durch rote Schriftfarbe betont. Die Untersuchung sollte aus Sicht die-ser möglichen Vorlieben bzw. Abneigungen vertieft werden. Die entsprechende Unter-suchung mit der **Binomialverteilung**³⁴⁵ (s.a. [Pkt. 4.3.1.3.](#)) findet sich in einem eige-nen Bereich jeweils unterhalb der kompletten Häufigkeitsauswertung (s. *das Beispiel un-ter [Pkt. 4.3.2.2.3.](#)*).
- ➔ Die SRP-Vorlage Stochastik bietet eine **zweistufige Eingrenzungsmöglichkeit** der untersuchten Daten, denn manipulative Eingriffe in Tageskasseneinnahmen erfolgen nicht (zwingend) ständig und entwickeln sich weiter, so dass Spuren wegfallen können.



Digitale Ziffernanalyse

Steuerpflichtige/r:

Steuernummer:

Vertiefungsuntersuchung nach Wochentagen:

▲
▼

Mo

³⁴³ Bei der Einerstelle werden Werte unter € 100 aus der Analyse ausgeschlossen, weil bei diesen die Einerstelle auf die beiden führenden Zahlenpositionen mit abweichender Benford-Verteilung fällt (s. [Pkt. 4.3.2.2.1.](#)).

³⁴⁴ Bei unterdrückten Ergebnissen des χ^2 -Tests (Chi-Quadrat-Test) ist die formatseitige Hervorhebung auffälliger Musterabweichungen deaktiviert.

³⁴⁵ Mit der Vertiefung einer Ziffernanalyse auf eine bestimmte Ziffer ändern sich die Beobachtungs- und Erwartungsverteilung in eine (dichotome) ja-nein-Gegenüberstellung „Ziffer x“ bzw. „alle anderen Ziffern“, für die die Binomialverteilung die einschlägige Beschreibung der Wahrscheinlichkeitsverhältnisse bildet. Die Idealwahrscheinlichkeiten richten sich dann nach der jeweilig zutreffenden Punkterwartung mit z.B. „10 %“ zu „90 %“ bei der Gleichverteilung.

Die Eingrenzungsoptionen für Jahre und **Wochentage** auf dem Tabellenblatt „digZA“ mittels Drehfeld sind für den schnellen Überblick gedacht und wirken nur auf die betreffende Ziffernanalyse. Auf diese Weise kann schnell nachvollzogen werden, ob sich Ziffernauffälligkeiten in Teildatenbeständen (offenbar) gegenseitig auflösen oder ob eine bestimmte Kombination zeitlicher Einschränkungen zu einem intensiven Manipulationsverdacht führt. Zur weiteren Abklärung einer solchen Manipulationsspur können die untersuchten Daten auf dem Tabellenblatt „Daten“ für alle Analysen eingeschränkt werden.

Im nebenstehenden Beispiel soll dazu hinterfragt werden, ob die

Einschränkung auf Kalenderjahr(e):	<input type="text"/>	<input type="text"/>
Teilzeitraum (zB Saison):	<input type="text" value="1"/> <input type="text" value="5"/> bis <input type="text" value="30"/> <input type="text" value="9"/>	
Wochentag(e):	<input type="text" value="5"/> <input type="text" value="6"/>	<input type="text"/> <input type="text"/> <input type="text"/>

Freitage und Sonnabende innerhalb der Sommersaison „1.5. bis 30.9.“ besonders auffällige Ziffernmuster aufweisen, denn es ist aus manipulationspsychologischer Sicht denkbar, dass an erfahrungsgemäß einnahmestärkeren Tagen eher verkürzt worden ist. Auf diese Weise können z.B. auch Unterschlagungen an bestimmten Wochentagen von Einnahmenhinterziehungen unterschieden werden. Einschränkungen auf dem Tabellenblatt „Daten“, die für alle Analysen der SRP-Vorlage Stochastik gelten, werden auf jedem Auswertungsblatt mit einem Hinweis wie in diesem Fall „Untersuchte Daten: Zeitraum 1.5. – 30.9. Wochentag(e) 5 6“ angezeigt.

- ➔ Zusätzlich kann über eine Ankreuz-Kästchenspalte links neben den hineinkopierten Daten auf dem Tabellenblatt „Daten“ ein **Ausschluss einzelner Datensätze** vorgenommen werden. Auch diese Abwahl wirkt auf die kompletten Untersuchungen der SRP-Vorlage Stochastik und wird auf den Auswertungsblättern mit dem Hinweis „Untersuchte Daten: Einzeldatenabwahl“ vermerkt. Die Funktion ist im Besonderen für die gemeinsame Sachaufklärung³⁴⁶ geeignet, wenn z.B. bei entdeckten Ziffernauffälligkeiten einzelne Tageskasseneinnahmen ausgeschlossen werden sollen, weil sie zu pauschal abgerechneten Veranstaltungen gehören. Über den Autofilter zur Abwahlspalte kann der jeweilige Stand der gemeinsamen Aufklärung komfortabel dokumentiert werden.
- ➔ Die ausdruckvorbereiteten Auswertungsblätter der SRP-Vorlage Stochastik zu den Ziffernanalysen enthalten Tabellen mit den beobachteten Ziffernhäufigkeiten und den gegenübergestellten erwarteten gem. Gleich- bzw. Benford-Verteilung, alle (Zwischen-) Ergebnisse des jeweiligen χ^2 -Tests (Chi-Quadrat-Test), die dazugehörigen Grafiken mit den verbildlichte Mustervergleichen sowie – bei Aktivierung – die Auswahl einer Ein-

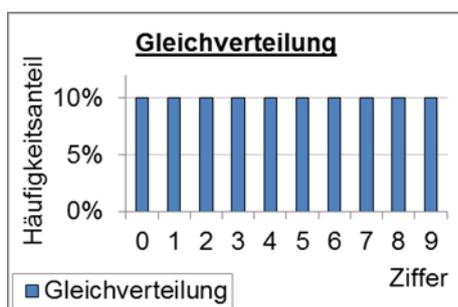
³⁴⁶ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit [...] Dabei ist allerdings die sphärenorientierte Mitwirkungspflicht des Stpfl. (s. § 88 Rz. 3 f.; § 90 Rz. 20) zu beachten. Mathematisch eindeutige Unschlüssigkeiten der Buchführungsergebnisse muss er durch betriebspezifische Besonderheiten schlüssig erklären“.

zelziffer mit den erforderlichen Parametern für und dem Ergebnis der Binomialverteilung als Übereinstimmungsprüfung. Damit werden alle Informationen ausgegeben, um die Zifferanalysen mit allgemein zugänglichen Mitteln reproduzieren zu können. Lediglich die verwendeten Daten müssen dazu z.B. im jeweiligen Bemerkungsfeld benannt werden.

4.3.2.2. Einerstellenuntersuchung

Bei der Einerstellenuntersuchung wird die Ziffernverteilung auf der Position vor dem Währungskomma in den zu untersuchenden Daten auf Übereinstimmung mit der „**Gleichverteilung**“ aller (möglichen) Ziffern geprüft.

4.3.2.2.1. Die (Punkt-)Erwartung „Gleichverteilung“



Die Annahme einer absolut zufälligen (gleichverteilten) Einerstelle insbesondere in Tageskasseneinnahmen gründet sich auf die **mathematisch theoretische** Überlegung, dass der Nachfragezufall „Anzahl an verkauften Artikeln bzw. Leistungen“ durch die Multiplikation mit den jeweiligen Einzelpreisen nicht beeinflusst

wird. Im Ergebnis würden sich die Ziffernverteilungen in den Tagessummen der Kasseneinnahmen genauso zufällig zusammensetzen wie die täglichen Nachfragehäufigkeiten.

Die **empirischen**³⁴⁷ **Erkenntnisse** des Benford-Gesetzes³⁴⁸ bestätigen die Gleichverteilung der Ziffern annähernd ab der Drittziffer mit zunehmender Genauigkeit zu den weiteren endständigen Positionen hin (s. [Pkt. 4.3.2.3.1.](#)). Aufgrund der abweichenden Erfahrungen für die ersten beiden Zahlenstellen gem. Benford-Gesetz werden in der SRP-Vorlage Stochastik Einzeldatensätze, die kleiner als € 100 sind, automatisch für die Einerstellenuntersuchung ausgeschlossen, weil deren **Einerstelle** auf einer der **ersten beiden Zahlenpositionen** liegt.

1.276, 20 *Tageskasseneinnahme*
 → ↑ **Blickrichtung**

tersuchung ausgeschlossen, weil deren **Einerstelle** auf einer der **ersten beiden Zahlenpositionen** liegt.

Obwohl die empirischen Beobachtungen zur Drittziffernverteilung gem. Benford-Gesetz eine Maximalabweichung der „0“ von knapp 1,8 % gegenüber der Gleichverteilung aufwei-

³⁴⁷ Empirische Erkenntnisse stammen aus Studien, die den Untersuchungsgegenstand unter nachvollziehbaren konstanten Bedingungen beobachten, und sollen theoretische Annahmen entweder bestätigen oder verwerfen. Gesicherte empirische Erkenntnisse genießen einen höheren Stellenwert in der Wissenschaft als widerstreitende theoretische Annahmen, weil sie die Realität beschreiben.

³⁴⁸ Benford, „The Law of Anomalous Numbers“, Proceedings of the American Philosophical Society 1938, 551.

sen, ist eine Anwendung letzterer auch auf dreistellige³⁴⁹ Tageskasseneinnahmen im Gesamtzusammenhang sinnvoll und sachgerecht:

-  Die Manipulationsform „frei erfundene Zahlen“, welche mit Ziffernanalysen aufgedeckt werden kann, kommt erfahrungsgemäß bei kleineren Bargeschäftebetrieben mit Tageskasseneinnahmen unter € 1.000 am ehesten vor.
- ⊙ Der maximale Fehler von 1,8 % zur erwarteten Gleichverteilung muss mit zwei anderen Größen im Kontext bewertet werden: Erstens wird eine Einerstellenauffälligkeit als Einzelindiz erst ab einer Nichtübereinstimmungswahrscheinlichkeit von 99,95 % aufgegriffen. Demgegenüber gilt zweitens das gesetzliche Beweismaß von ca. 95 % für die Beanstandung gem. § 158 AO (s. [Pkt. 4.6.](#)). Zudem treten derart eindeutige Ziffernauffälligkeiten nach ständigen Prüfungserfahrungen nur im Verbund mit gravierenden Ordnungsmängeln auf (s. [Pkt. 4.4.](#)).
- ☑ Darüber hinaus werden i.R. der SRP Ziffernauffälligkeiten so konkret herausgearbeitet (s. *das Beispiel unter* [Pkt. 4.3.2.2.3.](#)), dass Erfassungsfehler und einzelbetriebliche Sonderfaktoren als andere Ursachen zügig und mit geringem Aufwand abgeklärt werden können³⁵⁰.

Die Übereinstimmungsprüfung des beobachteten Häufigkeitsmusters aller Ziffern auf der Einerstelle der untersuchten Tageskasseneinnahmen mit der Gleichverteilung erfolgt durch den χ^2 -Test (Chi-Quadrat-Test). Dadurch beinhaltet das Übereinstimmungsergebnis die Datenanzahl (s. [Pkt. 4.3.1.3.](#)). Die übliche Empfehlung zur Mindestdatenmenge von 50 Datensätzen bei der Erwartung „Gleichverteilung“ aller Ziffern wird von der SRP-Vorlage Stochastik automatisch beachtet, und es werden die zwei Auffälligkeitsschwellen „99 %“ und „99,95 %“ formatseitig hervorgehoben (s. [Pkt. 4.3.2.1.](#)).

4.3.2.2.2. Nicht-manipulative Ursachen für Einerstellenauffälligkeiten

Einzelbetriebliche Besonderheiten als Ursache für mathematisch eindeutige (signifikante) Ziffernauffälligkeiten der Einerstelle können nur aus den Ursprungsgrößen der Tageskasseneinnahmen „Nachfragehäufigkeit(en)“ und „Einzelpreis(e)“ stammen und sind:

1. Praxisuntypische **besondere Einzelpreise** wie bspw. (fast) ausschließlich auf „0“ oder „5“ endende ohne Nachkommastellen können in der Multiplikation mit beliebigen Häufigkeiten ebenso nur die Einerstellen „0“ und „5“ erzeugen. Diese Ausnahme lässt sich anhand der verbildlichen Einerstellenuntersuchung sofort erkennen und durch Preislisten bestätigen.

³⁴⁹ Gemeint sind Tageskasseneinnahmen im Bereich „≥ € 100 bis € 999,99“.

³⁵⁰ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Dabei ist allerdings die sphärenorientierte Mitwirkungspflicht des Stpfl. (s. § 88 Rz. 3 f.; § 90 Rz. 20) zu beachten. Mathematisch eindeutige Unschlüssigkeiten der Buchführungsergebnisse muss er durch betriebsspezifische Besonderheiten schlüssig erklären“.

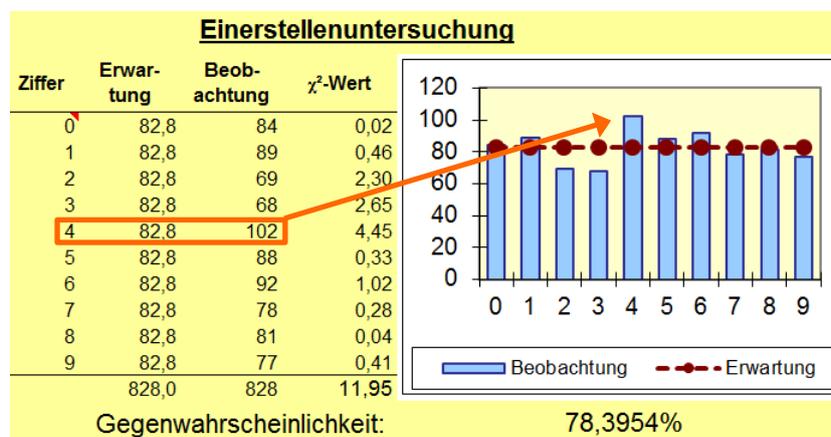
- Bei einer **täglich geringen Anzahl von verkauften Artikeln bzw. Leistungen** findet sich eine besondere Preisgestaltung in den Tagessummen tlw. wieder, wenn z.B. sehr wenige hochpreisige Produkte, welche alle auf „9“ enden, die Tageskasseneinnahmen bilden, werden die Endziffern „9“, „8“, „7“ etc. auf Kosten der kleineren gehäuft vorkommen. Diese Besonderheit lässt sich ebenfalls am dennoch gleichmäßigen Einerstellenmuster umgehend erkennen und durch die dann zwingend erforderliche Einzelaufzeichnung nachvollziehen.
- Tagesveranstaltungen, die über **Pauschalpreise** abgerechnet werden, entstehen nicht durch den Nachfragezufall sondern durch Preisverhandlungen. Befinden sich derartige Sondervorgänge unter den untersuchten Daten, können sie problemlos durch Filtern und Ankreuzen auf dem Tabellenblatt „Daten“ der SRP-Vorlage Stochastik nachvollziehbar aus allen Analysen ausgeschlossen werden.

Die Untersuchung der Einzelvorgänge aus Kassendaten auf die Gleichverteilung aller Ziffern an der Einerstelle ist nicht zielführend, weil der kombinatorische Zufall „zufällige Mengen multipliziert mit Einzelpreisen“ nicht wirksam ist, sondern die im Unternehmen festgelegte Preisstruktur die Ziffernverteilung bestimmt.

4.3.2.2.3. Schrittweise Abklärung von Einerstellenauffälligkeiten

Nach der **SRP-Systematik** erfolgt eine schrittweise Abklärung von Einerstellenauffälligkeiten. Einerseits können dadurch betriebliche Besonderheiten des Einzelfalls als Verursacher erkannt bzw. (weitest möglich) ausgeschlossen werden. Andererseits können partielle Manipulationen, wenn z.B. nur an erlösstärkeren Wochenendtagen eingegriffen wurde, aufgedeckt oder zwischen Einnahmenverkürzung und Unterschlagung unterschieden werden.

Diese beispielhafte Einerstellenverteilung zu den Tageskasseneinnahmen eines Eiscafés offenbart klare Unterschiede zwischen den Beobachtungsbalken und der erwarteten Gleichverteilung. Trotzdem ergibt der χ^2 -Test (Chi-Quadrat-Test) nur eine Nichtübereinstimmungswahrscheinlichkeit von ca. 78 % in Ab-



fés offenbart klare Unterschiede zwischen den Beobachtungsbalken und der erwarteten Gleichverteilung. Trotzdem ergibt der χ^2 -Test (Chi-Quadrat-Test) nur eine Nichtübereinstimmungswahrscheinlichkeit von ca. 78 % in Ab-

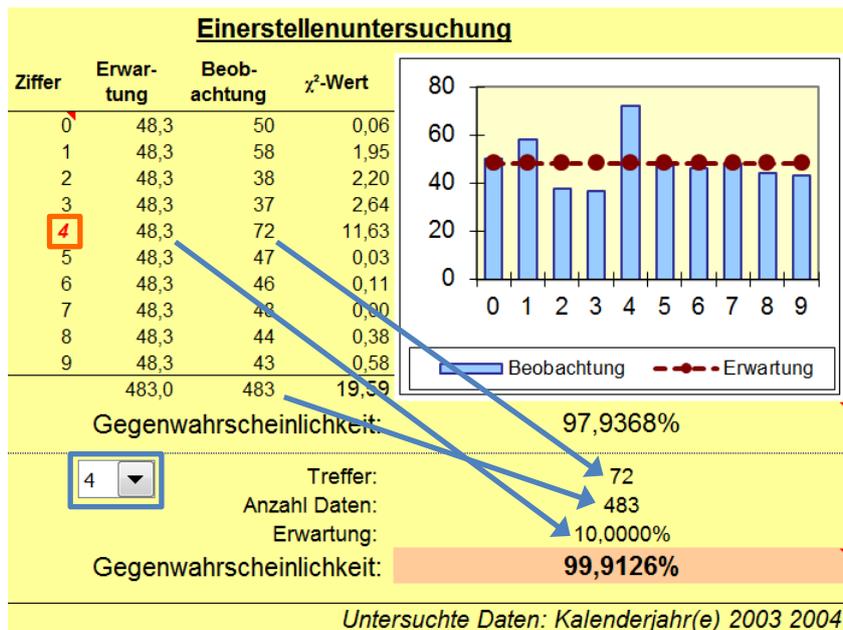
hängigkeit von der untersuchten Datenmenge „828 Tage“. Damit sind die beiden indiziellen Auffälligkeitsschwellen „99 %“ und „99,95 %“ (s. [Pkt. 4.3.2.1.](#)) nicht erreicht. Im verbildlichten

Mustervergleich fällt allerdings die **Häufung der „4“** auf, was sich als schwache Spur für eine Ziffernvorliebe in frei erfundenen Besteuerungsdaten herausstellen könnte.

Um diese Beobachtung besser zu verstehen, werden zunächst die drei Prüfungsjahre über das Auswahldrehfeld auf dem Tabellenblatt „digZA“ der SRP-Vorlage Stochastik einzeln betrachtet (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)). Dadurch lässt sich zügig feststellen, dass die Häufung der „4“ lediglich in den ersten beiden Jahren auftritt, im letzten nicht mehr. Gerade bei einfachen Eingriffsarten wie dem freien Erfinden von Zahlen ist das plötzliche Verschwinden von Manipulationsspuren nicht ungewöhnlich und kann eine Folge der Umstellung der bisherigen Hinterziehungsweise sein.

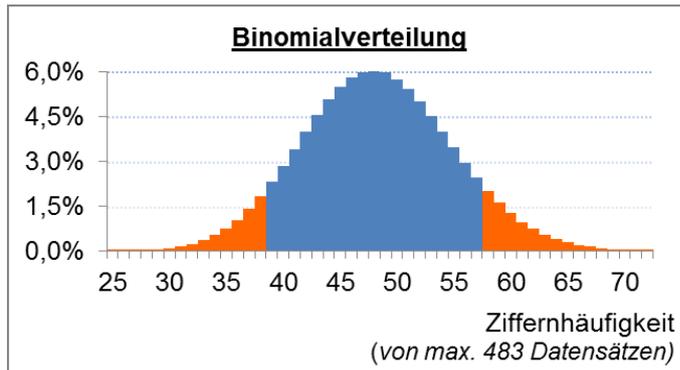
Deshalb wird die Einerstellenuntersuchung durch generelle Einschränkung der Daten auf dem Tabellenblatt „Daten“ auf die ersten beiden Jahre (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) begrenzt. In der Konsequenz gibt der χ^2 -Test (Chi-Quadrat-Test) nun eine Nichtübereinstimmungswahrscheinlichkeit von ca. 98 % aus. Damit werden die beiden Auffälligkeitsschwellen zwar weiterhin leicht unterschritten, die **Ziffer „4“** wird jetzt aber in der Ta-

belle mit den Häufigkeitsverteilungen **rot betont**. Diese Verfärbung weist darauf hin, dass eine Einzelbetrachtung der „4“-Häufung als mögliche Ziffernvorliebe sinnvoll ist. Dazu wird die „4“ im unterhalb befindlichen **Auswahldrehfeld des Bereichs „Einzeltest“** ausgewählt. Die Binomialverteilung wird akti-



viert und **übernimmt** die **Beobachtung** und **Erwartung** aus der Tabelle mit den Häufigkeitsverteilungen sowie die **Anzahl untersuchter Daten** und berechnet auf der Grundlage die Wahrscheinlichkeit, dass eine Einzelziffer so außergewöhnlich häufig oder selten vorkommen kann. Im Beispielsfall resultiert eine Gewissheit von ca. 99,9 %, dass diese Häufung nicht ohne menschlichen Einfluss geschehen konnte. Die rote Feldfarbe zeigt an, dass die erste Auffälligkeitsschwelle überschritten ist. Zusätzlich erscheint zur Auswertung der Hinweis auf die Einschränkung der untersuchten Daten, damit die Ziffernanalyse nachvollziehbar bleibt.

Für die Vertiefung der SRP-Ziffernanalysen auf eine Einzelziffer wird die **Binomialverteilung** in der vorsichtigsten Form angewendet. Das bedeutet die Gegenwahrscheinlichkeit einer häufig oder selten beobachteten Ziffer wird aus der Summe aller Einzelwahrscheinlichkeiten errechnet, dass eine Ziffer so häufig, noch häufiger, spiegelbildlich so selten oder noch seltener auftritt. Die nebenstehende Abb. gibt das Beispiel der Ziffer „1“ wieder. Da sie in den Tageskasseneinnahmen an der Einerstelle 58-mal gezählt wurde, werden die **Wahr-**



scheinlichkeitsbalken für „58-mal“, „59-mal“, „60-mal“ bis „483-mal“ sowie für „38-mal“, „37-mal“, „36-mal“ bis „0-mal“ **zusammengerechnet**. Die Summe beschreibt, wie wahrscheinlich es insgesamt ist, dass sich eine Ziffer so „außergewöhnlich“ zur (Punkt-)Erwartung verhält. Im Umkehrschluss drückt

die Gegenwahrscheinlichkeit maximal vorsichtig aus, ob eine Ziffer so fern der Erwartung liegen kann, ohne dass es sich um eine Ziffernvorliebe bzw. -abneigung handelt.

Abschließend ist für den konkreten Beispielsfall mit zu ca. 99,9 % unzufälliger Häufung der „4“ – **vor indizielle Beanstandung als Manipulationsspur** – noch zu klären, ob Erfassungsfehler oder eine der drei betrieblichen Besonderheiten (s. [Pkt. 4.3.2.2.2.](#)) zu der Zifferauffälligkeit geführt haben können. Preisgestaltungen können ausgeschlossen werden, weil selbst die simulierende Annahme eines Eiscafés mit nur einem Produkt zu € 4 an der Einerstelle der Tageskasseneinnahmen alle geraden Ziffern (gleichmäßig) aufweisen müsste. Auch eine zu geringe Anzahl an Einzelverkäufen ist bei einem Eiscafé nicht denkbar. Sofern sich unter den untersuchten Tagen nachweisbar Einzeltage mit Pauschalpreisen befinden, lassen sich diese zügig und ohne großen Aufwand auf dem Tabellenblatt „Daten“ der SRP-Vorlage Stochastik aus den Analysen entfernen. Im Ergebnis verbleibt nach der schrittweisen Abklärung der beobachteten Einerstellenauffälligkeit ein weitest gehend gegen andere Ursachen abgesicherter Manipulationsverdacht auf frei erfundene Tageskasseneinnahmen.

Die Beweiskraft von mathematisch eindeutigen (signifikanten) Einerstellenauffälligkeiten wird wegen der vergleichbaren Voraussetzungen aller manipulationsorientierten Musteranalysen in einem gemeinsamen Kapitel *unter* [Pkt. 4.3.5.](#) behandelt. Die derzeitige Beschränkung auf den Status eines „Zusatzindizes“³⁵¹ lässt sich durch verbreitete fehlerhafte Metho-

³⁵¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. Trossen; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. Zimmermann [Fahrten-

denbeschreibungen³⁵² und die resultierenden Zweifel an der korrekten Anwendung³⁵³ erklären.

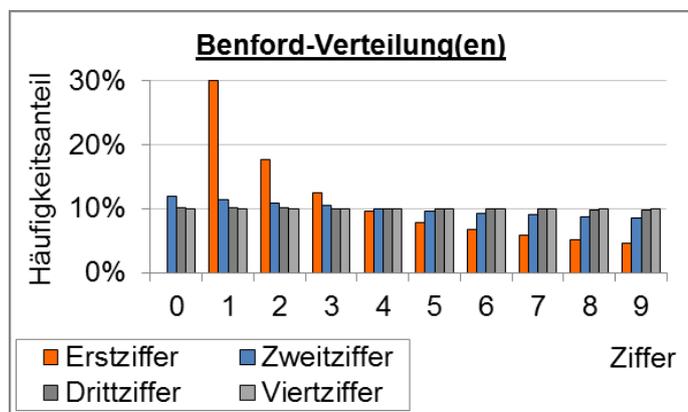
4.3.2.3. Zweitziffernanalyse (Benford-Gesetz)

Bei der Zweitziffernanalyse wird die **zweite schwebende**³⁵⁴ **Zahlenposition** in den zu untersuchenden Daten auf Übereinstimmung mit der Verteilung gem. **Benford-Gesetz** geprüft.

1 . 2 7 6 , 2 0 *Tageskasseneinnahme*
 → **Blickrichtung**

4.3.2.3.1. Die (Punkt-)Erwartung „Benford-Gesetz“

Empirische³⁵⁵ **Grundlage** für das Benford-Gesetz sind die übereinstimmenden Beobachtungen des Astronoms und Mathematikers Newcomb (1835-1909) sowie des Elektroingenieurs und Physikers Benford (1883-1948)³⁵⁶. Danach zeigen die führenden Positionen (besonders) in wirt-



buch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. Kühnen; Sosna, St-Bp. 2000, 41 [Teil I], 68 [Teil II]; Meyer, DStR 2005, 2114; Watrin/Stuffert, DB 2006, 1748; Huber, StBp. 2009, 121 [124 ff.].
³⁵² Vgl. z.B.: Bisle, „Chi-Quadrat-Test‘ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PStR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Erkenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.
³⁵³ Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“
³⁵⁴ In einem Datenbestand, der über einen Dezimalsprung hinausgeht wie bspw. „€ 800 bis € 1.500“, hat die Zweitziffer unterschiedliche Wertbedeutungen: bei „€ 800“ die Zehnerstelle und bei „€ 1.500“ die Hunderterposition.
³⁵⁵ Empirische Erkenntnisse stammen aus Studien, die den Untersuchungsgegenstand unter nachvollziehbaren konstanten Bedingungen beobachten, und sollen theoretische Annahmen entweder bestätigen oder verwerfen. Gesicherte empirische Erkenntnisse genießen einen höheren Stellenwert in der Wissenschaft als widerstreitende theoretische Annahmen, weil sie die Realität beschreiben.
³⁵⁶ „Newcomb-Benford-Law (NBL)“ bzw. „Benford-Gesetz“, veröffentlicht in: Benford, „The Law of Anomalous Numbers“, Proceedings of the American Philosophical Society 1938, 551.

schaftsnahen Daten charakteristische, von der Gleichverteilung abweichende Ziffernhäufigkeiten³⁵⁷: Für die **Erstziffer** ergibt sich – neben dem Fehlen der „0“ – eine starke Betonung der kleinen Ziffern auf zunehmende Kosten der höheren. Die **Zweitziffer** nähert sich bereits weitgehend der Gleichverteilung an.

Mathematisch theoretisch wird diese regelmäßige Erfahrung insbesondere mit der Dynamik der Verzinsung erklärt: Periodisch verzinstes Vermögen verweilt in jedem Dezimalbereich jeweils längere Zeit mit einer führenden „1“ als mit einer „2“, „3“, „4“ etc. Dadurch erhöht sich entsprechend die Wahrscheinlichkeit, Werte mit einer „1“ als Erstziffer zufällig herauszugreifen, gegenüber denen mit größeren Ziffern. Eine allgemeinere Erklärung findet sich im Zusammenhang zwischen Benford-Gesetz und Logarithmischer Normalverteilung³⁵⁸: Die Benford-Verteilungen beschreiben innerhalb der Logarithmischen Normalverteilung, welche in Wirtschaftsdaten überwiegend beobachtet wird (s. [Pkt. 4.3.4.1.](#)), die Ziffernmuster der Zahlenpositionen unter der Blickrichtung „von vorne“. Auf einer entsprechenden Verbindung zwischen Ziffernhäufigkeiten und (klassifizierter) Werteverteilung basieren auch die Ansätze der Wissenschaft, mit einer modifizierten Benford-Verteilung die Prüfung von Wirtschaftsdaten auf menschliche Eingriffe zu verbessern.³⁵⁹

Auf die Untersuchung der Erstziffern wird in der SRP von Beginn an verzichtet.³⁶⁰ Diese sind von – manipulationspsychologisch typischen (s. [Pkt. 4.3.4.](#)) – stärkeren Eingriffen in höhere Tageskasseneinnahmen zwar am ehesten betroffen, aber das Risiko von Fehldeutungen einer Ziffernanalyse ist ebenfalls am größten. Da sich die Tageserlöse eines einzelnen Unternehmens, in welchem frei erfundene Besteuerungszahlen erfahrungsgemäß vorkommen können, i.d.R. in einem begrenzten Zahlenbereich (auch) unterhalb € 1.000 bewegen, kann die Nachfragehäufung zur Betonung bestimmter Einzelziffern führen: So bewirkt

³⁵⁷ Das Benford-Gesetz wird durch eine allgemeine Berechnungsformel beschrieben: $P_{(D)} = \log_{10} \left(1 + \frac{1}{D} \right)$ für die Ziffer „D“.

³⁵⁸ Petersen, „Benford’s Law und die Betriebsprüfung – Eine Erwiderung auf Wähnerts „Interaktive Außenprüfung – Möglichkeiten der gemeinsamen Sachverhaltsaufklärung im digitalen Zeitalter“ (Stbg 2014 S. 20 ff.) aus wissenschaftshistorischer Perspektive“, Stbg 12/2015, 506, 510: „Ist die Kongruenz der diskreten Benford-Verteilung mit der stetigen logarithmischen Normalverteilung wohl in der Tat existent [...]“ mit Verweis auf: Scott/Fasli, „Benford’s law: An Empirical Investigation and a Novel Explanation“, CSM Technical Report 329, Essex 2001, 15.

³⁵⁹ Vgl. z.B. die digitale Spurensuche mit modifizierter Benford-Verteilung am Fraunhofer-Institut für Sichere Informationstechnologie auf der Internetseite der Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V.: „Digitale Spurensuche – Wirtschaftsthemen Mai 2011“, nachzulesen unter: <https://www.fraunhofer.de/de/presse/presseinformationen/2011/mai/it-forensik-gegen-wirtschaftskriminalitaet.html>.

³⁶⁰ Insofern verwundern umfangreiche Ausführungen zur Erstziffernverteilung gem. des Benford-Gesetzes i.R. einer Kritik an der SRP in: Petersen, „Benford’s Law und die Betriebsprüfung – Eine Erwiderung auf Wähnerts „Interaktive Außenprüfung – Möglichkeiten der gemeinsamen Sachverhaltsaufklärung im digitalen Zeitalter“ (Stbg 2014 S. 20 ff.) aus wissenschaftshistorischer Perspektive“, Stbg 12/2015, 506, 507 (Kapitel „Genesis von Benford’s Law“).

ein typischer Bereich an Tageskasseneinnahmen von „€ 500 bis € 800“ häufig die Erstziffern „5“, „6“ und „7“. Die Erstziffernkurve (s. *Abb. oberhalb*) verschiebt sich auf der Ziffernachse. Dieser störende Einfluss auf die Erstziffernanalyse tritt bei der Zweitzeiffer nur sehr begrenzt auf. Im Fall einer Zweitzeifferauffälligkeit wird i.R. der SRP-Ziffernanalytik die mögliche Verursachung durch einen begrenzten Zahlenbereich gezielt abgeklärt (s. *das Beispiel unter Pkt. 4.3.2.3.3.*).

Ein größenabhängiger Ausschluss von Daten wie bei den Werten unter € 100 i.R. der Einerstellenuntersuchung (s. *Pkt. 4.3.2.2.1.*) geschieht bei der Zweitzeifferanalyse in der SRP-Vorlage Stochastik nicht, weil die besonderen Ziffernmuster der führenden Zahlenpositionen durch die Veränderungen an den Dezimalsprüngen entstehen. In der Folge kann es dazu kommen, dass auf dem Tabellenblatt „digZA“ unterschiedlich viele Tageskasseneinnahmen in den beiden Ziffernanalysen der Einerstelle und Zweitzeiffer ausgewertet werden.

Die Übereinstimmungsprüfung des beobachteten Häufigkeitsmusters aller Ziffern an der zweiten Position der untersuchten Tageskasseneinnahmen mit der Benford-Verteilung erfolgt durch den χ^2 -Test (Chi-Quadrat-Test). Dadurch beinhaltet das Übereinstimmungsergebnis die Datenanzahl (s. *Pkt. 4.3.1.3.*). Die übliche Empfehlung zur Mindestdatenmenge von 59 Datensätzen³⁶¹ bei der Erwartung „Benford-Verteilung für die Zweitzeiffer“ wird von der SRP-Vorlage Stochastik automatisch beachtet, und es werden die zwei Auffälligkeitsschwellen „99 %“ und „99,95 %“ formatseitig hervorgehoben (s. *Pkt. 4.3.2.1.*).

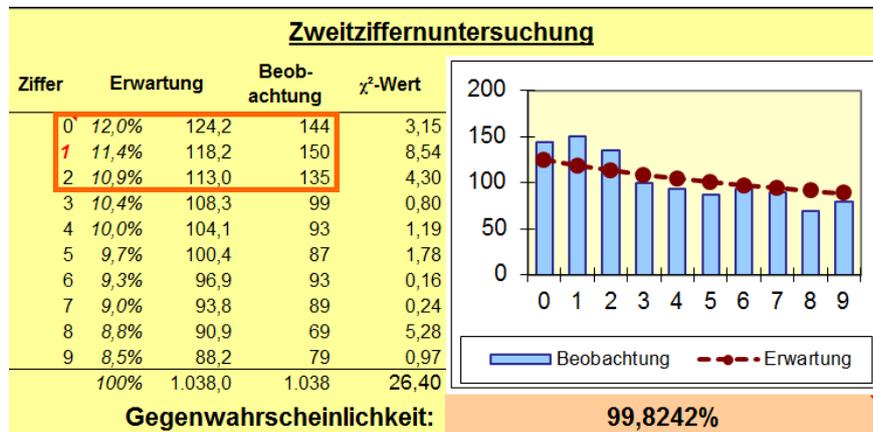
4.3.2.3.2. Nicht-manipulative Ursachen für Zweitzeifferauffälligkeiten

Einzelbetriebliche Besonderheiten als Ursache für mathematisch eindeutige (signifikante) Ziffernauffälligkeiten der Zweitposition haben im Gegensatz zur Einerstelle i.d.R. einen Bezug zu den Größenverhältnissen der untersuchten Zahlen: Sofern die betreffenden Tageskasseneinnahmen entweder an einem **Dezimalsprung** oder in einem **begrenzten Bereich oberhalb von € 1.000** ihre Häufungszone haben (s. *Pkt. 4.3.2.3.3.*), werden dadurch bestimmte Zweitzeiffern betont: Ein unternehmenstypischer Schwerpunkt der Tageskasseneinnahmen von „€ 2.200 bis € 2.600“ bringt überwiegend die Zweitzeiffern „2“, „3“, „4“ und „5“ hervor. Derartige Konstellationen sind eher selten, weil bei einem höheren Einnahmenniveau auch die Streuung meistens größer ist. Dennoch wird auch diese mögliche praxisnahe Besonderheit des Einzelbetriebs i.R. der SRP-Ziffernanalytik gezielt abgeklärt, wobei die Funktionalitäten der SRP-Vorlage Stochastik helfen.

³⁶¹ Die Mindestdatenmenge von 59 Datensätzen ergibt sich aus dem Prozentanteil von 8,5 % für die nach dem Benford-Gesetz am seltensten vorkommende „9“ angewendet auf die Empfehlung „ ≥ 5 erwartete Elemente pro Klasse“.

4.3.2.3.3. Schrittweise Abklärung von Zweitziffernauffälligkeiten

Nach der **SRP-Systematik** erfolgt eine schrittweise Abklärung von Zweitziffernauffälligkeiten, um betriebliche Besonderheiten des Einzelfalls als Verursacher erkennen bzw. (weitestgehend) ausschließen zu können.



Die exemplarisch abgebildete Häufigkeitsverteilung zur Zweitziffer in den Tageskasseneinnahmen eines gastronomischen Betriebs stimmt nicht mit der (Punkt-)Erwartung gem. Benford-Gesetz

überein. Der χ^2 -Test (Chi-Quadrat-Test) bestätigt den visuellen Eindruck mit einer sehr hohen Nichtübereinstimmungswahrscheinlichkeit von ca. 99,8 % oberhalb der ersten Auffälligkeitsschwelle (s. das Stichwort „Chi-Quadrat-Test“ unter [Pkt. 4.3.2.1.](#)). Im Widerspruch zum Einsatzzweck der Ziffernanalyse, Vorlieben, Abneigungen oder Chaos aufzudecken, ist allerdings eine recht gleichmäßige **Überbetonung der kleinen Ziffern „0“, „1“ und „2“** auf Kosten fast aller größeren abzulesen. Dabei weist die Rotfärbung der Ziffer „1“ darauf hin, dass sie aus der Einzelperspektive ebenfalls mathematisch eindeutig (signifikant) die Erwartung verfehlt (s. das Beispiel im Bereich „Binomialverteilung“ unter diesem Pkt.). **Bei zwar eindeutigen aber in dieser Form gleichmäßigen Zweitziffernauffälligkeiten muss die Ursache einer Bündelung der untersuchten Zahlen an einem Dezimalsprung abgeklärt werden.**

Einen ersten Hinweis auf den Streubereich der untersuchten Tageskasseneinnahmen

statistische Kenngrößen	
absolute Kenngrößen (lineare Schichtung)	
Anzahl:	1.038
<i>Anzahl ab Schichtungsminimum:</i>	
Summe:	884.561,93
Mittelwert:	852,18
Standardabweichung:	371,93

men geben die statistischen Kenngrößen auf dem Tabellenblatt „Daten“ in der SRP-Vorlage Stochastik. Mit einem Mittelwert von € 852,18 und

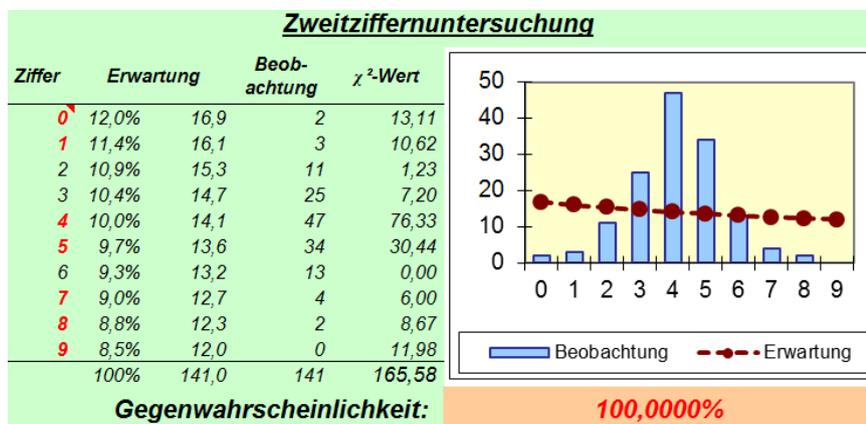
einer vorzeichenneutralen durchschnittlichen Abweichung jedes einzelnen Datensatzes zum Mittelwert (**Standardabweichung**) von € 371,93 reicht der Normalbereich der Tageskasseneinnahmen über den Dezimalsprung „Hunderter auf Tausender“ hinaus: Mittelwert + Standardabweichung = € 1.224,11. Wie stark gehäuft die Daten zwischen € 1.000 und € 1.224,11 vorkommen, lässt sich aus den statistischen Kenngrößen nicht direkt ablesen, weil die Häufigkeitsverteilung „Logarithmische Normalverteilung“ asymmetrisch ist (s. [Pkt. 4.3.4.1.](#)).

Weil die betriebliche Besonderheit „Häufung der Tageskasseneinnahmen an einem Dezimalsprung“ als Ursache nicht ausgeschlossen werden kann, muss die Zweitziiffernanalyse weiter vertieft werden. Dazu bieten die Klassifizierungen der Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ der SRP-Vorlage Stochastik schnelle und zielführende Funktionen mit ihren Unterteilungen der Daten in gleichmäßige Werteklassen. Die weitere Beschreibung dieser Untersuchung eines auffälligen Zweitziiffernergebnisses auf eine mögliche nicht-manipulative Ursache erfolgt in dem betreffenden Abschnitt „Abklärung von Zweitziiffernauflälligkeiten“ unter [Pkt. 4.3.3.2.](#)

Als weiteres Beispiel für die schrittweise Abklärung eines mathematisch eindeutigen (signifikanten) Zweitziiffernmusters haben die Tageskasseneinnahmen eines Fischgeschäfts insgesamt ein zu häufiges Vorkommen der Ziffern „4“ und „5“ ergeben. Eine zufällige Abweichung von der Benford-Verteilung kann bei einer Nichtübereinstimmungswahrscheinlichkeit gem. χ^2 -Test (Chi-Quadrat-Test) von mehr als 99,99 % ausgeschlossen werden. Die Suche nach auslösenden Teilen in den Daten mit Hilfe der Auswahl-Drehfelder für Jahre und Wochentage auf dem Tabellenblatt „digZA“ (s. den Bereich „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) ergibt für mehrere Wochentage signifikant vom Benford-Gesetz abweichende Häufigkeitsmuster, die aber dabei gleichmäßig aussehen.

Als Ursprung der insgesamt zu häufigen „4“ und „5“ fällt umgehend der Donnerstag mit der abgebildeten Zweitziiffernverteilung auf.

Trotz der mathematisch eindeutigen (signifikanten) Nichtübereinstimmung mit der Erwartung gem. Benford-Gesetz und diverser rot hervorgehobener (s. das Bei-



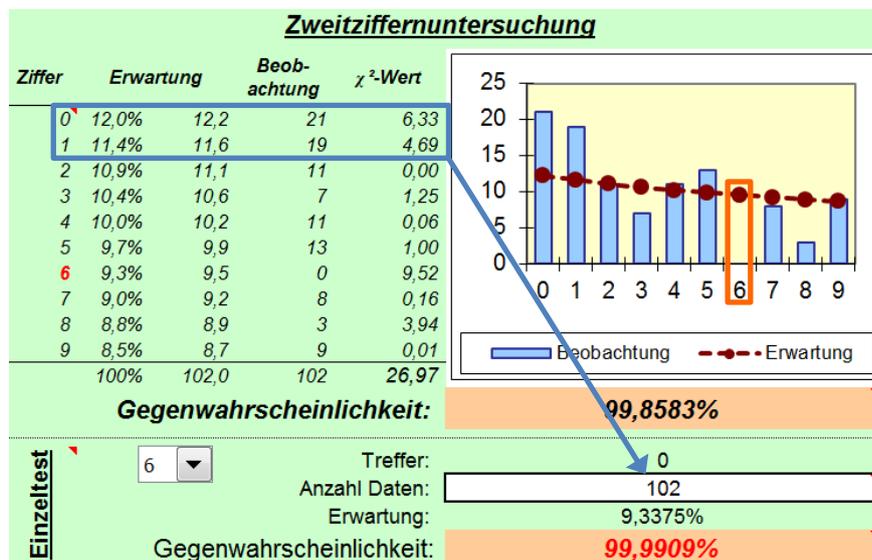
spiel im Bereich „Binomialverteilung“ unter diesem Pkt.) Ziffern deutet die glockenförmig gleichmäßige Häufigkeitskurve nicht auf bestimmte Ziffernvorlieben oder -abneigungen bzw. -chaos hin. Vielmehr ist dieses Verteilungsmuster von Zweitziiffern in Tageskasseneinnahmen typisch für Werte, die in einem begrenzten Zahlenbereich oberhalb von € 1.000 liegen. Weil die täglich aufsummierten Bareinnahmen eines einzelnen Betriebs i.d.R. nicht allzu weit auf der Zahlenskala streuen, wirkt die Verzehnfachung³⁶² der Trefferwahrscheinlichkeit für

³⁶² Während die Spanne bspw. für die Zweitziiffer unterhalb des Dezimalsprungs „Hunderter auf Tausender“ mit z.B. „€ 950 bis € 959,99“ ein Intervall von € 10 umfasst, wächst sie oberhalb auf € 100 an: „€ 1.500 bis € 1.599,99“. Die Wahrscheinlichkeit, dass eine Tageskasseneinnahme in die jeweiligen Bereiche zufällig hineinfällt, verändert sich vergleichbar.

Zweitziffern nach dem Dezimalsprung „Hunderter auf Tausender“ nicht auf alle Ziffern gleichermaßen. Für das abgebildete Beispiel würde eine Konzentration der Tageskasseneinnahmen im Bereich „€ 1.300 bis € 1.599,99“ in Betracht kommen. Auch die Abklärung dieser möglichen Ursache für die beobachtete Zweitziffernauffälligkeit gelingt mit Hilfe der Klassifizierungen der Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ in der SRP-Vorlage Stochastik und wird *unter Pkt. 4.3.3.2.* dargestellt.

Ein Perspektivwechsel auf eine Einzelziffer, die an der Zweitposition besonders häufig oder selten beobachtet wurde, ist auf dem Tabellenblatt „digZA“ der SRP-Vorlage Stochastik auch vorgesehen. In Betracht kommende Ziffern werden in der Tabelle mit den Häufigkeitsverteilungen durch rote Schrift hervorgehoben. Sie können im Bereich „Einzeltest“ unterhalb über ein Drehfeld ausgewählt werden. Die Informationen „Treffer“ und „Trefferwahrscheinlichkeit“³⁶³ zu dieser Einzelziffer sowie Datenanzahl, welche für die **Binomialverteilung** als einschlägigem Übereinstimmungstest erforderlich sind, werden übernommen und die Nichtübereinstimmungswahrscheinlichkeit ausgegeben.

Eine Besonderheit ist bei der Einzelziffernbetrachtung i.R. der Zweitziffernanalyse gegenüber der Einerstellenuntersuchung zu beachten: Wie aus der Grafik erkennbar verschlimmern neben dem



verschlimmern neben dem auffälligen Fehlen der Ziffer „6“ auch die zu häufigen Ziffern „0“ und „1“ das Ergebnis des χ^2 -Tests (Chi-Quadrat-Test) entscheidend. Mit Hilfe der Werteklassifizierungen der SRP-Vorlage Stochastik kann zünftig herausgefunden werden, dass diese Ta-

gesskasseneinnahmen eines Eiscafés am Dezimalsprung „Hunderter auf Tausender“ liegen (s. *Pkt. 4.3.3.2.*) und deshalb vermehrt Nullen und Einsen an der Zweitposition aufweisen. Dieser relevanteste Störeinfluss für die Zweitziffernanalyse wirkt sich auch auf die Einzelziffernbetrachtung aus, weil die Datensätze, welche durch den Dezimalsprung an „0“ und „1“ gebunden sind, nicht zu den erwarteten 9,5-mal „6“ beitragen können. Ausschließlich für

³⁶³ Im Gegensatz zur Einerstellenuntersuchung sind bei der Zweitziffernanalyse die Trefferwahrscheinlichkeiten der Einzelziffern nicht gleich, sondern müssen aus der Benford-Verteilung übernommen werden. So ist z.B. die Wahrscheinlichkeit für eine „0“ an der Zweitposition ca. 12 %. Somit ergibt sich als Summenwahrscheinlichkeit für alle anderen Ziffern der Gegenwert zu 100 %, also ca. 88 %.

diese Konstellation besteht in der SRP-Vorlage Stochastik die Möglichkeit, die Datenanzahl für die Binomialverteilung anzupassen³⁶⁴. Auf diese Weise kann die Wahrscheinlichkeit für das auffällige Fehlen der „6“ trotz Überlagerung mit einem Dezimalsprung abgeschätzt werden. Der Kommentar zum Eingabefeld benennt diesen besonderen Anwendungsfall der Zweitziiffernanalyse.

Die Beweiskraft von mathematisch eindeutigen (signifikanten) Zweitziiffernauffälligkeiten wird wegen der vergleichbaren Voraussetzungen aller manipulationsorientierten Musteranalysen in einem gemeinsamen Kapitel *unter Pkt. 4.3.5.* behandelt. Die derzeitige Beschränkung auf den Status eines „Zusatzindizes“³⁶⁵ lässt sich durch verbreitete fehlerhafte Methodenbeschreibungen³⁶⁶ und die resultierenden Zweifel an der korrekten Anwendung³⁶⁷ erklären.

4.3.3. Klassifizierung (Schichtung) der Werte

Bei der Klassifizierung (Schichtung³⁶⁸) der zu untersuchenden Daten werden diese in Werteklassen unterteilt, um durch anschließende Verbildlichung die Häufigkeitsmuster herauszuarbeiten. Dazu ist i.d.R. die Festlegung von gleichgroßen³⁶⁹ Klassenabständen zielführend.

³⁶⁴ Im Beispielsfall weist die Tabelle mit den Häufigkeitsverteilungen zur Zweitziiffer aus, dass die Ziffern „0“ und „1“ $(21+19) - (12,2+11,6) = 16,2$ -mal zu oft vorkommen. Als Annäherung kann diese Häufigkeit dem Dezimalsprung angelastet werden, so dass die verfügbare Datenanzahl für eine „6“ auf $102 - 16,2 =$ abgerundet 85 korrigiert wird.

³⁶⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. Trossen; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. Zimmermann [Fahrtenbuch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. Kühnen; Sosna, StBp. 2000, 41 [Teil I], 68 [Teil II]; Meyer, DStR 2005, 2114; Watrin/Stuffert, DB 2006, 1748; Huber, StBp. 2009, 121 [124 ff.]“.

³⁶⁶ Vgl. z.B.: Bisle, „Chi-Quadrat-Test“ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PStR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Erkenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.

³⁶⁷ Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

³⁶⁸ Für die Einteilung in Größenklassen werden unterschiedliche Begriffe verwendet wie bspw. „Schichtung“ im Datenanalyseprogramm IDEA® oder „Gruppierung“ in MS Excel®.

³⁶⁹ Sofern ein logarithmisches Muster wie die Logarithmische Normalverteilung erwartet wird, ist die Wahl von Größenklassen mit logarithmisch gleichen Abständen (z.B. „ $10^2 - 10^3 - 10^4 - 10^5$ etc.“ entsprechend „100 –

I.R. der SRP-Systematik (s. den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)) haben die dafür vorgesehenen Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ keine eigenständige Prüffunktion, sondern dienen der **Orientierung in großen Datenmengen** (s. das Beispiel im Abschnitt „gleichmäßige Muster der Häufigkeiten“ unter [Pkt. 4.3.](#)) sowie der **Absicherung manipulationsorientierter Musteranalysen**:

- © Das **Suchen nach prüfungsrelevanten Vorgängen in unprüfbar großen Datenmengen** anhand von Häufigkeitsmustern lehnt sich an eine zentrale Erkenntnis der Datenanalytik an: „Komplexität hat sehr viel mit Vernetzung zu tun, ja kommt erst durch Vernetzung zustande. Komplexe Vorgänge verlangen daher zu ihrem Verständnis ein Denken in Zusammenhängen, das sich an der Struktur organisierter Systeme und ihrer speziellen Dynamik orientiert. [...] Damit treffen wir auf einen der Kardinalfehler in der Beurteilung, was für die Erfassung von Komplexität wichtig ist. Denn hier geht es um das ‚Gesicht‘ der Wirklichkeit, um das Erkennen von Mustern“³⁷⁰ (s.a. das Beispiel zur Eingrenzung einer Zeitreihenauffälligkeit im Abschnitt „Prüfungsnetz“ unter [Pkt. 2.3.](#)). Da betriebsuntypische Sachverhalte häufig aus den gleichmäßigen Häufigkeitsmustern ausbrechen, bietet sich auf diesem Weg mit Hilfe digitaler Lösungen eine effektive Möglichkeit, diejenigen Vorgänge effizient zu identifizieren, die ein besonders hohes Fehler- oder Manipulationsrisiko in sich tragen. Dazu verwendet die SRP-Vorlage Stochastik auf dem Tabellenblatt „Schichtung“ den externen Vergleich mit dem allgemeinen Häufigkeitsmuster in Wirtschaftsdaten „Logarithmische Normalverteilung“ (s. [Pkt. 4.3.4.1.](#)). Zusätzlich werden auf dem Tabellenblatt „Schichtungsvertiefung“ bis zu drei Jahre und Wochentage nach freier Auswahl über Drehfelder intern gegenübergestellt. Weitere Funktionalitäten der SRP-Vorlage Stochastik wie bspw. die Umstellung von linearer auf logarithmische Klassifizierung (s. den Abschnitt „Umstellung von linearer auf logarithmische Klassenbildung“ unter [Pkt. 4.3.3.1.](#)) unterstützen die variable Musteranalyse für ein zielgenaues und zügiges Auffinden von betriebsspezifischen Prüffeldern.
- © Für möglichst eindeutige Indizien aus den beiden manipulationsorientierten Musteranalysen der SRP durch (weitest gehende) Abklärung anderer denkbarer Erklärungen werden **Zweitziffernauffälligkeiten** mittels besonderer Klassifizierung **abgesichert** (s. die beiden Beispiele im Abschnitt „Abklärung von Zweitziffernauffälligkeiten“ unter [Pkt. 4.3.3.2.](#)) und die **Verteilungsanalyse** (s. [Pkt. 4.3.4.](#)) **vorbereitet**. Damit kann

1.000 – 10.000 – 100.000 usw.“) oft zielführender, um die Gleichmäßigkeit der beobachteten Verteilung visuell zu beurteilen.

³⁷⁰ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 16 u. 21.

bei der Zweitziffernanalyse die praxisrelevanteste nicht-manipulative Ursache für Auffälligkeiten bestätigt oder ausgeschlossen werden. Für die Untersuchung von Tageskasseneinnahmen auf – manipulationspsychologisch typische (s. [Pkt. 4.3.4.](#)) – stärkere Eingriffe in die höheren Werte wird bereits im Vorfeld der wichtigste Störfaktor „uneinheitlicher Datenbestand“ abgeklärt. Im Ergebnis erreichen danach verbleibende Musterauffälligkeiten eine größere indizielle Gewissheit (s. [Pkt. 4.3.5.](#)).

4.3.3.1. Anwendungshinweise zu SRP-Werteklassifizierungen

Folgende Hinweise zur Durchführung von Datenklassifizierungen mit der SRP-Vorlage Stochastik sind wichtig:

➔ Alle auf dem Tabellenblatt „Daten“ vorgenommenen **generellen Dateneinschränkungen** (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) wirken auch auf die Werteklassifizierungen der SRP-Vorlage Stochastik durch. Ein entsprechender Hinweis erfolgt auf den Auswertungsblättern „Schichtung“ und „Schichtungsvertiefung“.

➔ Um mittels **Klassenunterteilung** Häufigkeitsmuster in den zu untersuchenden Daten erkennen zu können, muss zunächst eine zielführende Anzahl an Klassen festgelegt werden. Danach kann diese auf den relevanten Wertebereich angewendet werden, wodurch der einheitliche Klassenabstand berechnet wird. Infolge einer unvoreilhaften Klassenfestlegung verteilen sich die Daten entweder auf zu wenige Intervalle oder werden zu kleinteilig abge-

bildet, so dass jeweils das Streumuster schlecht sichtbar wird. Die SRP greift auf die prüfungserprobte **Wurzelmethode** zurück: Dabei ergibt sich die Klassenanzahl als [Wurzel aus der Datenanzahl](#), wird aber zur

statistische Kenngrößen	
<i>absolute Kenngrößen (lineare Schichtung)</i>	
Anzahl:	1.094
<i>Anzahl ab Schichtungsminimum:</i>	
Summe:	513.530,85
Mittelwert:	469,41
Standardabweichung:	260,27
Minimum:	63,30
Hilfsminimum:	140,00
Maximum:	2.369,40
Hilfsmaximum:	1.230,00
Vorschlag Intervallanzahl:	30
Vorschlag Schichtungsintervall:	40,00
Vorschlag erste Schichtobergrenze:	140,00
Vorschlag zweite Schichtobergrenze:	180,00

besseren Übersichtlichkeit [auf maximal 30 Intervalle begrenzt](#). Durch die Auflistung der statistischen Parameter neben den eingefügten Daten auf dem Tabellenblatt „Daten“ wird die automatische Klassenbildung nachvollziehbar dokumentiert. In dem Beispiel oberhalb ergäbe die hohe Datenanzahl mit $\sqrt{1.094} = \text{ca. } 33$ Intervalle, so dass in der SRP-Vorlage Stochastik die Beschränkung auf 30 Klassen vorgenommen wird. Diese Klassenanzahl wird auf den Wertebereich vom Minimum bzw. Hilfsminimum (s. den

nächsten Auflistungspkt. „Hilfsminimum und Hilfsmaximum“) bis zum Maximum bzw. Hilfsmaximum gerundet angewendet, wodurch sich die ersten beiden (und alle weiteren) Schichtobergrenzen sowie das feste Intervall gem. Abb. ergeben. Nach jeder Dateneinschränkung auf dem Tabellenblatt „Daten“ (s. die zweite Abb. zu „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) wird die automatische Klassifizierung umgehend automatisch angepasst. Neben der empfohlenen Klassenbildung können bei jeder Schichtung die ersten Obergrenzen (und damit die Intervallgröße) manuell abgeändert werden (s. den übernächsten Auflistungspkt. zur „freien Klassenbildung“).

- Die SRP-Vorlage Stochastik versucht für zielführendere Werteklassifizierungen extreme Einzelwerte ohne engeren Verteilungsbezug zu den untersuchten Daten (Ausreißer) selbständig zu erkennen und verwendet in dem Fall die Visualisierungsgrenzen „**Hilfsminimum**“ und „**Hilfsmaximum**“. Dazu werden die Erfahrungen nach *Grubbs*³⁷¹ abgewandelt herangezogen: Datensätze, welche mehr als den dreifachen Interquartilsabstand³⁷² nach oben oder unten vom Median³⁷³ abweichen, gelten als „Ausreißer“. Für die manipulationsorientierte Verteilungsanalyse (s. [Pkt. 4.3.4.](#)) sind die Grenzen mit dem fünffachen Interquartilsabstand weiter gefasst, weil hier das Muster wichtiger als das Hervortreten von außergewöhnlichen Einzelsachverhalten ist. Erkannte Ausreißer werden durch die Hilfswerte als Visualisierungsgrenzen der Klassenfestlegung ersetzt. Dies wird aus den „statistischen Kenngrößen“ auf dem Tabellenblatt „Daten“ wie in der Abb. zum vorigen Aufzählungspunkt ersichtlich, indem die Werte „Minimum“ und „Maximum“ hellgrau unterdrückt sind. Die betreffenden Ausreißer-Datensätze werden dadurch **nicht** aus den Analysen ausgeschlossen. Ist dies gewünscht, müssen sie wie nebenstehend **per Kreuz einzeln abgewählt** werden.

	Datum	alle Werte	Bemerkung	untersuchte Werte	lg
x	13.05.2006	2.360,40			
	14.05.2006	482,80		482,80	2,684
	15.05.2006	150,60		150,60	2,178
	16.05.2006	286,40		286,40	2,457

Daraufhin erscheint auf jedem Auswertungsblatt automatisch der Hinweis „Untersuchte Daten: Einzeldatenabwahl“.

- Alle Datenklassifizierungen der Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ können auch mit **freier Klassenbildung** durchgeführt werden. Dazu lassen sich die ersten zwei Schichtobergrenzen in den beiden freien Feldern oberhalb einzeln oder insgesamt abändern. Die SRP-Vorlage Stochastik führt das durch die Eingabe(n) entste-

³⁷¹ *Grubbs*, „Engineering Statistics Handbook“, Kapitel 7.1.6. „What are outliers in the data?“, nachzulesen unter: <https://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm>.

³⁷² Der Interquartilsabstand einer Datenmenge beschreibt den Wertumfang zwischen dem 25 %- und dem 75 %-Quantil (1. und 3. Quartil). Dazu wird die Differenz zwischen dem Wertinhalt des – aufsteigend sortiert – 75 %- und 25 %-größten Rangs ermittelt.

³⁷³ Der Median ist der mittlere, also der 50 %-Rang einer Datenmenge.

hende Intervall automatisch fort, bis der höchste Wert der untersuchten Daten bzw. die maximale Klassenanzahl „30“ auf dem Tabellenblatt „Schichtung“ resp. „28“ auf dem Tabellenblatt „Schichtungsvertiefung“ erreicht ist. Steht eine **manuell eingetragene Schichtobergrenze im Widerspruch zu einem eingetragenen Startwert** (s. den übernächsten Aufzählungspkt. zum „Startwert“), färbt sich die Schrift rot. Dieser Hinweis hat keinen Einfluss auf die gewählte Klassifizierung. Für die Gegenüberstel-

Schicht- obergrenze
300
200,00
400,00
200
400
600

lungen der Häufigkeitsmuster nach Jahren bzw. Wochentagen kann die (manuell abgeänderte) Klassenbildung vom Tabellenblatt „Schichtung“ mittels Anhak-Kästchen neben den freien Eingabefeldern für die ersten beiden Schichtobergrenzen auf dem Tabellenblatt „Schichtungsvertiefung“ übernommen werden.

- ➔ Falls infolge der (automatischen oder einer manuell abgeänderten) Klassenfestlegung **nicht alle hohen Spitzenwerte in den Auswertungen berücksichtigt** werden können, verfärbt sich das Summenfeld unter der jeweiligen Häufigkeitsverteilung auf den Tabellenblättern „Schichtung“ und „Schichtungsvertiefung“ rot.

- ➔ Sowohl für automatische als auch für manuell abgeänderte Klassenbildung bietet die SRP-Vorlage Stochastik auf den Tabellenblättern „Schichtung“ und „Schichtungsvertiefung“ jeweils **per Anhakfeld eine Umstellfunktion von linearer auf logarithmische Klassifizierung** an.

Schicht- obergrenze	Beob- achtung
	Abs
10,00	0
20,00	0
10	0
20	0
40	0
80	2
160	41

Logarithmische Klassen entwickeln sich über gleichmäßig zunehmende Exponenten: z.B. „ $10^1 - 10^2 - 10^3 - 10^4$ “ entsprechend „10 – 100 – 1.000 – 10.000“. Durch die zunehmend größer werdenden Intervalle lassen sich Klassifizierungen zu sehr hohen Werten ausweiten. Falls ein logarithmisches Häufigkeitsmuster wie die „Logarithmische Normalverteilung“ (s. [Pkt. 4.3.4.1.](#)) erwartet wird, erleichtert eine logarithmische Klassenbildung die visuelle Übereinstimmungsprüfung durch Symmetrierung.

- ➔ Auf dem Tabellenblatt „Schichtung“ können alle untersuchten Daten **per Anhakfeld als vorzeichenneutrale Beträge** verarbeitet werden. Diese Funktion betont häufige Korrekturbuchungen bzw. -aufzeichnungen in den

Schicht- obergrenze	Beob- achtung	LogNorm- Verteilung
		Absolutwerte: <input checked="" type="checkbox"/>
		logarithmische Klassifizierung
70	2	0
140	18	21
210	103	88
280	135	145

gleichmäßigen Häufigkeitsmustern, wodurch sie sich als Prüffelder leichter finden lassen. Die Aktivierung der Funktion wird über die statistischen Kenngrößen auf dem Ta-

ellenblatt „Daten“ sowie jeweils oberhalb der Häufigkeitsgegenüberstellungen auf dem Tabellenblatt „Schichtungsvertiefung“ durch Rotfärbung von „Minimum“ und „Maximum“ angezeigt. Zudem wird dann die Berechnung eines Hilfsminimums und -maximums unterdrückt (s. *den Auflistungspkt. „Hilfsminimum und Hilfsmaximum“ oberhalb*).

- ➔ Einheitlich für alle Werteklassifizierungen der Tabellenblätter „Schichtung“ und „Schichtungsvertiefung“ kann ein **Startwert** festgelegt werden, ab dem die musterbasierten Häufigkeitsanalysen ausgeführt werden. Hilfreich ist diese Funktion, um untypische kleine Werte auszuschließen oder den Untersuchungsfokus auf einen eigenständigen Teil hochwertiger Daten zu legen. Alle empfohlenen Schichtungsparameter (s. *den Auflistungspkt. „Klassenunterteilung mit Wurzelmethode“ oberhalb*) werden automatisch auf die neuen Verhältnisse umgestellt. Zu den „statistischen Kenngrößen“ auf dem Tabellenblatt „Daten“ wird die „Datenanzahl ab Schichtungsminimum“ angegeben. Auf dem Tabellenblatt „Schichtungsvertiefung“, auf welches der Startwert automatisch übertragen wird, zeigt die rote Feldfarbe zur „maximalen Datenanzahl“ jeweils über den Häufigkeitsauswertungen die manuell vorgenommene untere Begrenzung an.
- ➔ Für die Gegenüberstellungen der Häufigkeitsmuster der bis zu drei auswählbaren Jahre oder Wochentage auf dem Tabellenblatt „Schichtungsvertiefung“ werden die Datenreihen als **Liniengrafen** wiedergegeben, obwohl es sich um diskrete³⁷⁴ Verteilungen handelt. Ebenso wird das externe Vergleichsmuster „Logarithmische Normalverteilung“ auf dem Tabellenblatt „Schichtung“ als Linie und Fläche abgebildet. Beides erfolgt aus dem Grund, durch eine bessere Übersichtlichkeit die visuelle Übereinstimmung und außergewöhnliche Vorgänge leichter erkennen zu können.
- ➔ Keine der Klassifizierungen auf den Tabellenblättern „Schichtung“ und „Schichtungsvertiefung“ wird mit einem **Wahrscheinlichkeitstest** auf Übereinstimmung geprüft. Da die Auswertungen (lediglich) die Funktion „Orientierung in großen Datenumfängen“ und „Absicherung entdeckter Ziffernauffälligkeiten gegen nicht-manipulative Ursachen“ haben und nicht indiziell verwendet werden, ist ein qualifizierter Abschluss unter Einbeziehung der untersuchten Datenmenge (s. [Pkt. 4.3.1.3.](#)) nicht erforderlich. Herausstechende Einzelsachverhalte werden als betriebsspezifische Prüffelder nach dem Auffinden anhand der zugehörigen Unterlagen (rechtlich) beurteilt. Bei der Abklärung einer Zweitziffernauffälligkeit gibt die besondere Schichtungsform ausreichend Aufschluss darüber, ob ein begrenzter Zahlenbereich am Dezimalsprung „Hunderter auf Tausender“ oder oberhalb von € 1.000 ursächlich sein kann (s. *die Beispiele unter [Pkt. 4.3.3.2.](#)*). Sollen

³⁷⁴ Unter diskreten Verteilungen werden in der Mathematik solche verstanden, die keine Zwischenwerte einnehmen können. So kann es bei Münzwürfen aus Sicht einer bestimmten Münzseite nur zu ganzzahligen Erfolgen (bzw. Misserfolgen) kommen, nicht aber zu 3,437 Treffern. Das Gegenteil mit kontinuierlichen Ergebnismöglichkeiten wird „stetig“ genannt.

Tageskasseneinnahmen auf die – manipulationspsychologisch typischen (s. [Pkt. 4.3.4.](#)) – stärkeren Eingriffe in höhere Werte untersucht werden, dienen die Wertklassifizierungen als Vorbereitung zum Erkennen der praxisrelevanten Besonderheit „uneinheitliche Mehrfachverteilungen“ (s. *das Beispiel unter* [Pkt. 4.3.3.3.](#)). Die qualifizierte Übereinstimmungsprüfung mit einem Wahrscheinlichkeitstest zur Berücksichtigung der untersuchten Datenmenge erfolgt i.R. der Verteilungsanalyse (s. [Pkt. 4.3.4.2.](#)).

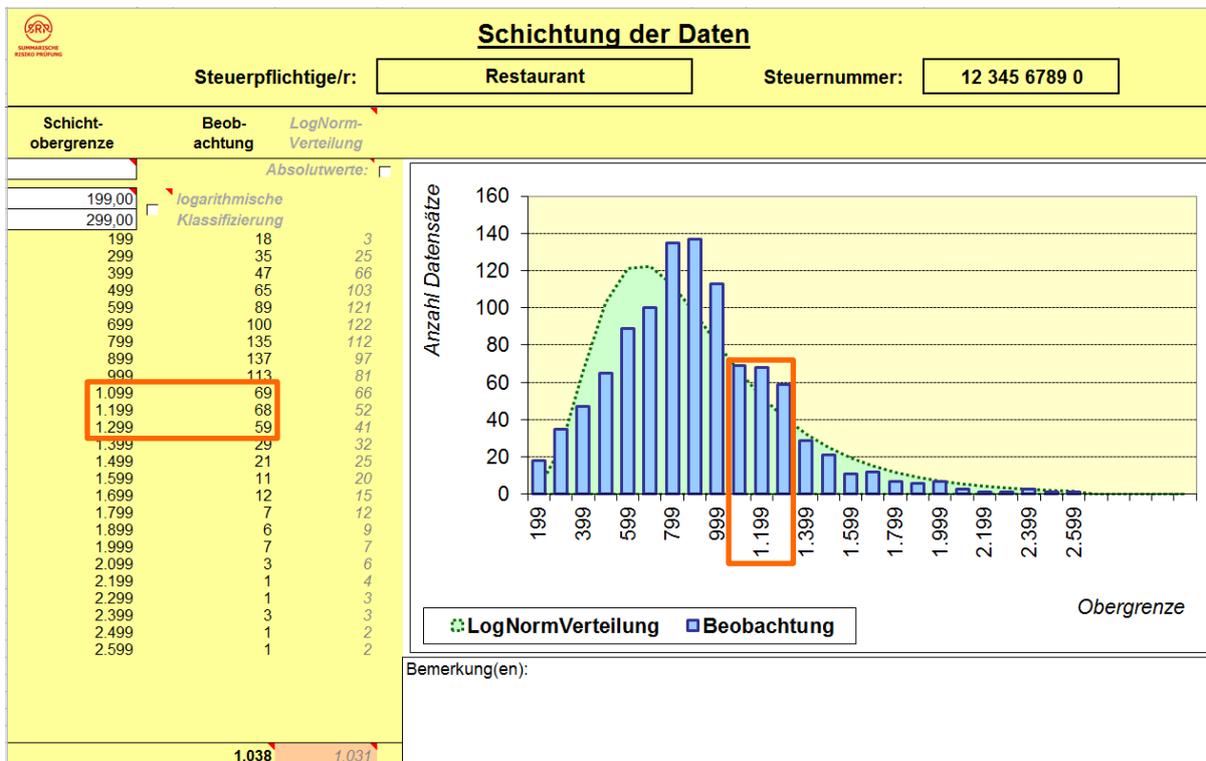
4.3.3.2. Abklärung von Zweitziiffernauffälligkeiten mit 99-er Schichtung

Aufgedeckte Ziffernauffälligkeiten haben nur dann indiziellen Aussagewert gegen die sachliche Richtigkeit der Besteuerungsgrundlagen, wenn die anderen beiden Formen menschlicher Eingriffe (s. *oberhalb des Abschnitts „musterbasierte Prüfungsansätze zur Aufdeckung von Manipulationen“ unter* [Pkt. 4.3.](#)) (weitest gehend) ausgeschlossen werden können. Während dies bei der Einerstellenuntersuchung anhand (ordnungsmäßiger) Betriebsunterlagen und (mathematisch) einfacher Überlegungen geschieht (s. [Pkt. 4.3.2.2.2.](#) sowie zum Ende des Beispiels unter [Pkt. 4.3.2.2.3.](#)), hilft bei der Abklärung von Zweitziiffernauffälligkeiten eine **besondere Form der Musteranalyse**: Durch die Überprüfung, ob eine Konzentration der untersuchten Tageskasseneinnahmen in einem begrenzten Zahlenbereich am Dezimalsprung „Hunderter auf Tausender“ oder oberhalb von € 1.000 vorliegt, lassen sich beide praxisrelevanten Sondergründe für auffällige Zweitziiffernmuster (s. [Pkt. 4.3.2.3.2.](#)) mittels **99-er Schichtung**³⁷⁵ einfach und schnell bestätigen bzw. ausschließen.

Das erste Beispiel mit auffälliger Zweitziiffernverteilung zu den Tageskasseneinnahmen eines gastronomischen Betriebs (s. [Pkt. 4.3.2.3.3.](#)) wies eine recht gleichmäßige Überbetonung der Ziffern „0“, „1“ und „2“ auf, was in der Gesamtbetrachtung der gegenübergestellten Ziffernverteilungen „Beobachtung“ und „(Punkt-)Erwartung“ gem. χ^2 -Test (Chi-Quadrat-Test) zu einer Nichtübereinstimmungswahrscheinlichkeit von ca. 99,8 % geführt hat. Die vorgefundene Konstellation mit den zu häufigen kleinsten Ziffern an der Zweitposition auf Kosten aller größeren ist das typische Bild eines **Dezimalsprungs**: Vor dem praxisrelevantesten Dezimalsprung in möglicherweise frei erfundenen Tageskasseneinnahmen „Hunderter auf Tausender“ umfasst die Zweitziiffernspanne € 10 wie bspw. in „€ 950 bis € 959,99“. Direkt danach wächst sie abrupt auf € 100 wie z.B. in „€ 1.500 bis € 1.599,99“ an. Die Wahrscheinlichkeit, dass ein Datensatz in die jeweiligen Bereiche zufällig hineinfällt, erhöht sich ent-

³⁷⁵ Exakt müsste für eine gezielte Zweitziiffernbeobachtung oberhalb von € 1.000 mit Schichtobergrenzen gearbeitet werden, die auf „99,99“ enden (z.B. „€ 1.100 bis € 1.199,99“). Bei einer entsprechenden Eingabe auf dem Tabellenblatt „Schichtung“ in der SRP-Vorlage Stochastik ist darauf zu achten, dass im Diagramm auf ganze Zahlen gerundet wird und somit jeweils die vorige Zweitziiffer „klassenbestimmend“ ist: Im Beispiel der Klasse „€ 1.100 bis € 1.199,99“ zeigt die X-Achse als Obergrenze „€ 1.200“ an, wobei alle Werte des Intervalls eine „1“ an der zweiten Position haben müssen.

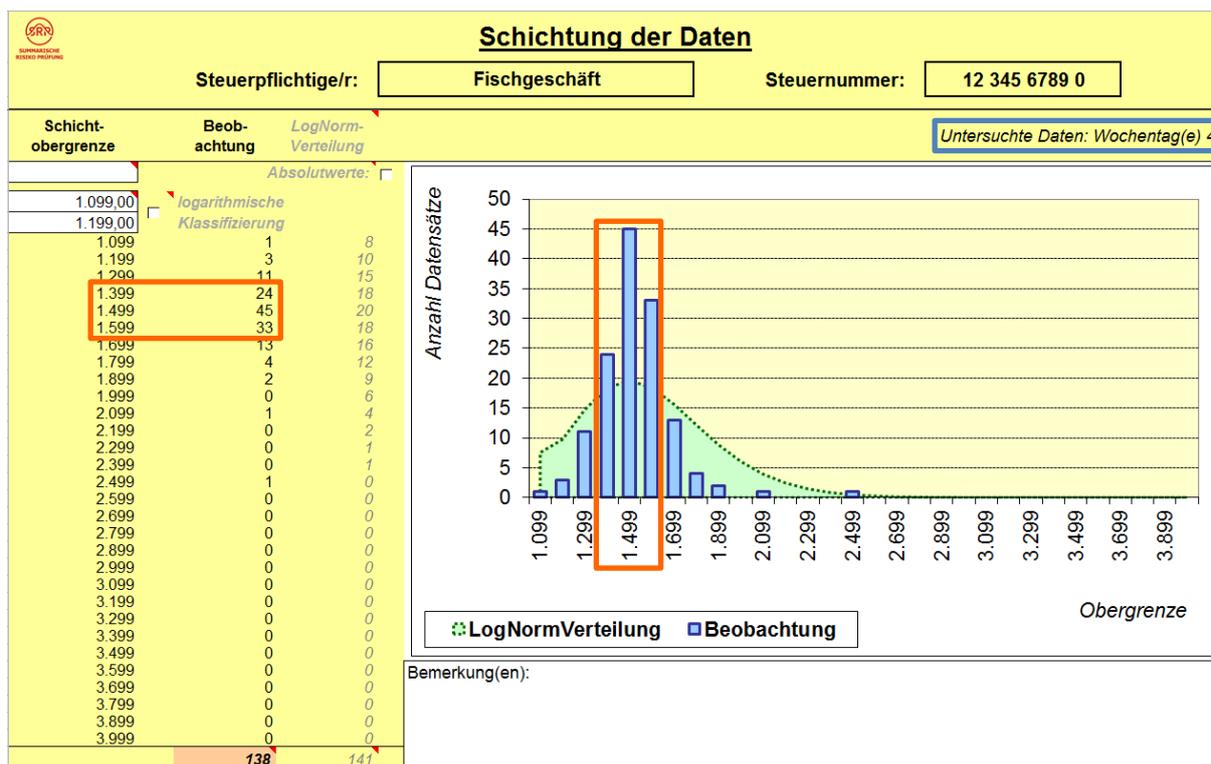
sprechend. Durch die besondere 99-er Schichtung der auffälligen Tageskasseneinnahmen, mit welcher die Zweitposition oberhalb von € 1.000 auf eine bestimmte Zweitziffer festgelegt wird, lässt sich diese nicht-manipulative Ursache des Einzelfalls umgehend bestätigen oder ausschließen:



Sowohl der Häufigkeitentabelle als auch dem Diagramm ist leicht zu entnehmen, dass mit **fast 200 Tageskasseneinnahmen** ein erheblicher Teil der Datensätze in den Zahlenbereich fällt, der **nach dem Dezimalsprung zwingend eine „0“, „1“ oder „2“ an der Zweitposition** haben muss. Da es sich hierbei nicht nur um ein paar Werte im Verteilungsauslauf handelt, ist somit zumindest eine Mitverursachung der Zweitzifferauffälligkeit bestätigt. **In der Konsequenz ist das mögliche Indiz gegen die sachliche Richtigkeit entkräftet.**

Im zweiten Beispielsfall konnte die auffällige Häufung der „4“ und „5“ an der Zweitstelle der Tageskasseneinnahmen eines Fischgeschäfts (s. [Pkt. 4.3.2.3.3.](#)) mit Hilfe der vertiefenden Auswahl-Drehfelder für Jahre und Wochentage des Tabellenblatts „digZA“ auf den Donnerstag als Hauptverursacher konkretisiert werden. Die glockenförmig gleichmäßige Abnahme der Zweitziffern von der „3“, „4“ und „5“ zu den kleineren und größeren hin passt aber ebenfalls nicht zu den erwarteten Manipulationsspuren frei erfundener Daten „Ziffernvorlieben, -abneigungen oder -chaos“. In solchen Fällen kann eine recht seltene Situation für die betreffenden kleineren Betrieben mit Bargeschäften vorliegen: Es wird **regelmäßig ein höheres Niveau der Tageskasseneinnahmen oberhalb von € 1.000** erreicht, **wobei** aber die **Streuung** aufgrund einer sehr gleichmäßigen Nachfrage durch vorwiegend

Stammkundschaft **begrenzt ausfällt**. Auch diese einzelbetriebliche Besonderheit kann als mögliche nicht-manipulative Ursache für eine beobachtete Zweitziiffernauffälligkeit mit der 99-er Schichtung problemlos abgeklärt werden. Im Beispielsfall ist dazu zunächst die grundsätzliche Einschränkung der untersuchten Daten auf Donnerstage im entsprechenden Bereich des Tabellenblatts „Daten“ vorzunehmen (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit unter [Pkt. 4.3.2.1.](#)):

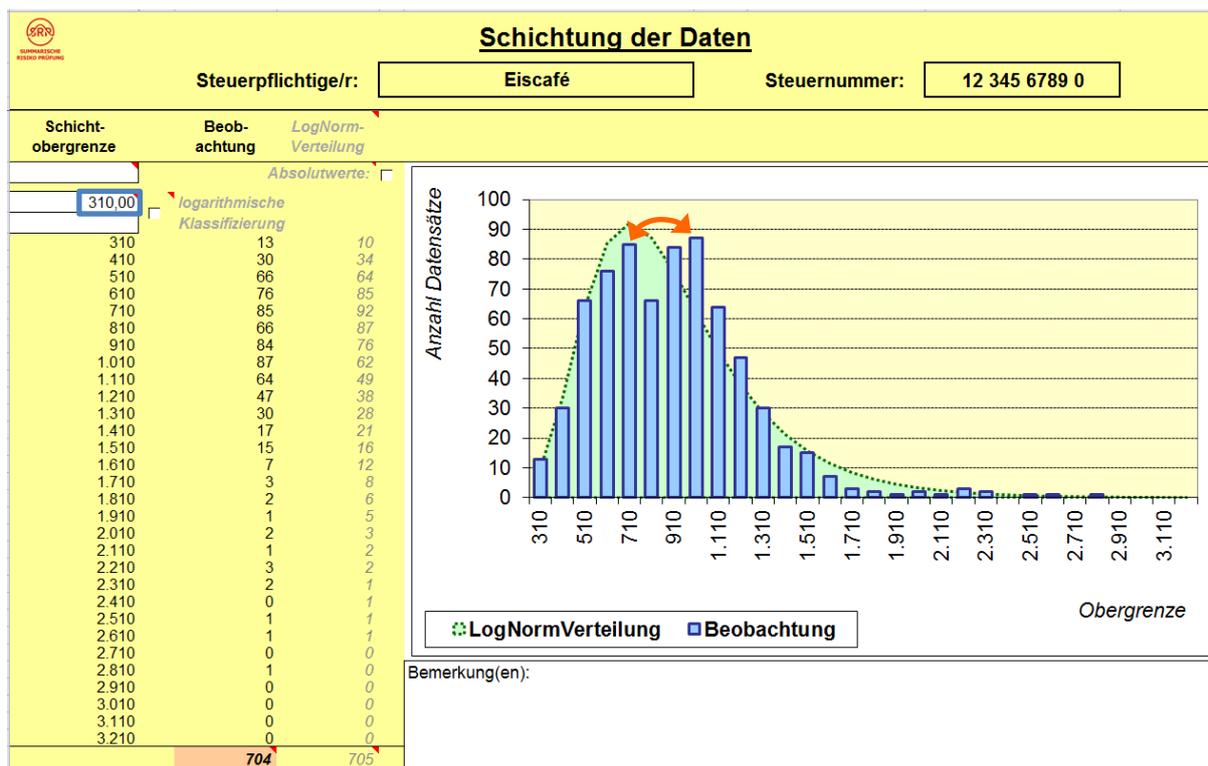


Die Annahme einer schmalbandigen Konzentration der täglichen Bareinnahmensummen aller Donnerstage oberhalb von € 1.000 mit den daraus hervorgehenden Ziffern „4“ und „5“ bestätigt sich durch die Häufungszone „ab € 1.300 bis € 1.599“. Davon haben sich die „4“ und „5“ als Abweichung gegenüber der Benford-Verteilung offensichtlich bis in den Gesamtdatenbestand aller Wochentage durchgesetzt. **Somit ist der Verdacht frei erfundener Tageskasseneinnahmen anhand der beobachteten Zweitziiffernauffälligkeit nach klassifizierender Vertiefung nicht aufrecht zu halten.** *Auf der anderen Seite ergibt sich die wertvolle Erkenntnis für betriebswirtschaftliche Verprobungen wie bspw. Kalkulationen, dass mit der gleichmäßigen Nachfrage über alle Jahreszeiten hinweg eine ausgesprochen gute Planungssituation bestand.* Die notwendige Einschränkung der Daten auf die Donnerstage wird auf dem Auswertungsblatt angezeigt.

4.3.3.3. Vorbereitende Wertklassifizierungen für die Verteilungsanalyse

I.R. der SRP-Verteilungsanalyse soll das klassifizierte Häufigkeitsmuster von Tageskasseneinnahmen mit der Logarithmischen Normalverteilung (s. [Pkt. 4.3.4.1.](#)) verglichen werden,

um die kriminalpsychologisch wahrscheinlichste Manipulationsspur „verstärkte Spitzeneingriffe in die höheren Werte“ aufzudecken (s. [Pkt. 4.3.4.](#)). Dazu wird aus den beiden Kennzahlen „Mittelwert“ und „Standardabweichung“³⁷⁶ der untersuchten Daten *eine* einzelfallspezifische Logarithmische Normalverteilung abgeleitet. Deren asymmetrisches Auslaufen in Zahlenbereiche, die deutlich über dem mittleren Niveau der Tageskasseneinnahmen liegen (s. [Pkt. 4.3.4.1.](#)), erfasst auch die (wenigen) außergewöhnlichen Nachfragetage eines Unternehmens. Dennoch kann in besonderen Fällen eine so **uneinheitliche Absatzsituation** vorliegen, dass bestimmte Zeiten wie bspw. Wochenendtage oder eine Hochsaison nicht zur eigentlichen Häufigkeitsverteilung passen. Daraus resultiert eine Überlagerung von *zwei* Verteilungen, welche nicht mit *einer* reproduzierten Absatzverteilung übereinstimmen können. Die praxisorientierten Gegenüberstellungsoptionen von Jahren und Wochentagen auf dem Tabellenblatt „Schichtungsverteilung“ in Verbindung mit den Dateneinschränkungsmöglichkeiten auf dem Tabellenblatt „Daten“ (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) ermöglichen ein effizientes Erkennen solcher betriebsuntypischen Zeiten und die anschließende separate Untersuchung.

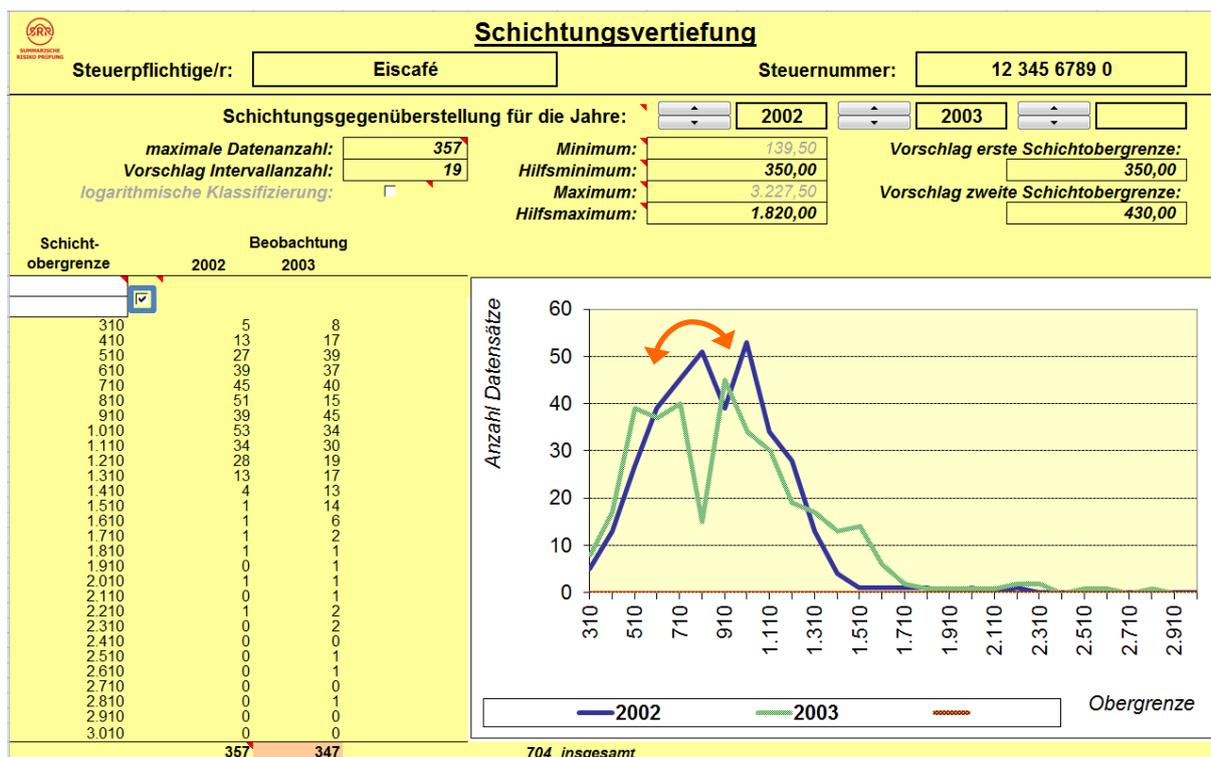


Im abgebildeten Beispiel zu den Tageskasseneinnahmen eines Eiscafés wurde die automatische Klassifizierung der SRP-Vorlage Stochastik durch **Abändern der ersten Schicht-**

³⁷⁶ Die Standardabweichung beschreibt die Streuungsintensität einer Datenmenge als vorzeichenneutrale durchschnittliche Abweichung jedes einzelnen Datensatzes zum Mittelwert.

obergrenze etwas größer³⁷⁷ durchgeführt, um die Einheitlichkeit des Verteilungsmusters besser erkennen zu können. Im visualisierten Ergebnis ist problemlos zu sehen, dass eine **doppelgipflige Überlagerung von zwei Häufigkeitsverteilungen** einer einheitlichen (Punkt-)Erwartung gegenübersteht. Folglich kann ein Übereinstimmungstest nur zu dem Urteil der Nichtübereinstimmung kommen.

Vor einer indiziellen Untersuchung der Tageskasseneinnahmen auf Spitzeneingriffe mit der Verteilungsanalyse (s. [Pkt. 4.3.4.](#)) müssen die beiden Absatzbereiche separiert werden. In einem ersten Schritt werden dazu auf dem Tabellenblatt „Schichtungsvertiefung“ die beiden Jahre des Prüfungszeitraums mit **aktivierter Übernahme der Schichtobergrenzen vom Tabellenblatt „Schichtung“** gegenübergestellt:



Diese Vertiefung löst das Problem nicht, denn beide Jahresgrafen weisen – neben anderen Unterschieden – **eine Unterteilung mit zwei Spitzen** auf.

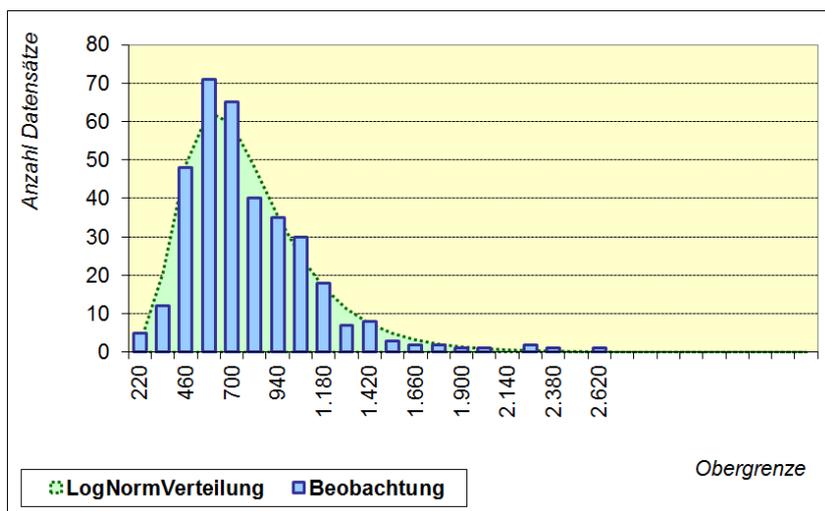
Der interne Vergleich der Häufigkeitsmuster für die einzelnen Wochentage auf dem Tabellenblatt „Schichtungsvertiefung“ legt offen, dass die Freitage und Sonnabende grundsätzlich ein höheres Einnahmenniveau erreichen³⁷⁸. Durch die Einschränkung der untersuchten Daten auf die Wochentage „1“, „2“, „3“, „4“ und „7“ bzw. „5“ und „6“ auf dem Tabellenblatt „Daten“ (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) kann

³⁷⁷ Wie bei der SRP-Zeitreihensystematik (s. den Abschnitt „Fuzzy Theorien“ unter Pkt. 4.2.2.) entscheidet die Intervallgröße bei Klassifizierungen darüber, ob entweder durch umfangreiche Klassen der grobe Blick auf das Häufigkeitsmuster oder durch kleinere das Erkennen von Ausreißern in den Fokus gelangt.

³⁷⁸ Die gesamten Liniengrafen der Häufigkeitsverteilungen zu den Freitagen und Sonnabenden liegen im Diagramm weiter rechts bei den höheren Werten auf der X-Achse als derjenigen der anderen Tage.

umgehend geprüft werden, ob diese Beobachtung ursächlich für die zwei Gipfel im Häufigkeitsmuster ist. Auch diese Differenzierung der Tageskasseneinnahmen führt jedoch nicht zu einer einheitlichen Verteilung.

Für die letzte praxistypische Option zur Aufteilung der untersuchten Daten in uneinheitliche Teilbestände wird mit Hilfe der Einschränkungsoptionen auf dem Tabellenblatt „Daten“ eine Haupt- und eine Nebensaison ausgewählt (s.a. das Beispiel im Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)). Im Beispielsfall löst sich die Überlagerung von zwei Verteilungen für die abgebildete – manipulationspsychologisch unkritischere – Neben-



saison „1.10. bis 31.3.“ weitest gehend auf. Demgegenüber ergibt sich für die Hauptsaison – selbst unter Einschränkung auf den Zeitraum „1.6. bis 30.9.“ – eine chaotische Häufigkeitsverteilung. Infolge der daraus resultierenden Schlussfolgerung, dass die Tageskassenein-

nahmen frei erfunden³⁷⁹ sein können, sind Wertklassifizierungen wie die SRP-Verteilungsanalyse mit Vorsicht auszuwerten, weil jeglicher Verteilungszusammenhang dadurch zerstört sein kann (s. den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)). Falls die Verteilungsanalyse zu keinem klaren Auffälligkeitsbild (s. [Pkt. 4.3.4.3.](#)) kommt, kann dem Verdacht eines unsystematischen Eingriffs in die höheren Saisoneinnahmen bspw. mit einer punktuellen Intensivierung der Formalprüfung (s. [Pkt. 4.1.](#)), einer Wetterzeitreihe für den betreffenden Zeitraum (s. den Abschnitt „Zeitreihenvergleiche mit externen Daten“ unter [Pkt. 4.2.10.](#)) oder anderen Gesamtbildmethoden (s. [Pkt. 4.4.](#)) nachgegangen werden.

4.3.4. Verteilungsanalyse

Die SRP-Verteilungsanalyse in der SRP-Vorlage Stochastik ist auf das Ziel ausgerichtet, speziell in Tageskasseneinnahmen **Spuren von Spitzeneingriffen** – also der verstärk-

³⁷⁹ Falls die Ziffernanalysen keinen mathematisch eindeutigen (signifikanten) Hinweis auf Spuren des freien Erfindens von Tageskasseneinnahmen ergeben haben, ist das dennoch keine Garantie für die gegenteilige Annahme. Als beispielhaftes Manipulationsmodell wird das regelmäßige Herausnehmen von glatten Hunderterbeträgen keine Ziffernvorlieben oder –abneigungen hinterlassen, kann das Häufigkeitsmuster jedoch bis zum Chaos zerstören.

ten Verkürzung höherer Werte – aufzudecken. Grundsätzlich können Vergleiche der klassifizierten Häufigkeitsverteilung von Besteuerungsdaten mit der Logarithmischen Normalverteilung (s. [Pkt. 4.3.4.1.](#)) – im Gegensatz zu Ziffernanalysen – unterschiedlichste Formen von menschlichen Eingriffen anzeigen³⁸⁰, i.R. der SRP wird die Anwendung aber auf die eine praxisrelevante Manipulationsart beschränkt, um das Risiko von Fehldeutungen der Musteranalysen zu minimieren.

Spitzeneingriffe sind auch unter der Verwendung moderner Manipulationshilfsmittel [s. den Abschnitt „raffinierte (digitale) Manipulationswerkzeuge unter [Pkt. 2.1.](#)] das **wahrscheinlichste aufdeckbare Eingriffsmuster**. Nach neurophysiologischen Erkenntnissen werden die allermeisten menschlichen Handlungen nicht bewusst gesteuert, sondern unterbewusst entschieden und im Nachhinein (möglichst) rational erklärt.³⁸¹ Dabei lösen Gesetzesbrüche zunächst eine – nicht unbedingt bewusst wahrgenommene – (adrenerge³⁸²) Stressreaktion aus. Tritt im Anschluss ein persönlicher Erfolg durch den Gesetzesbruch ein, wird zudem das Belohnungshormon „Dopamin“ ausgeschüttet. Bereits ein stark aktiviertes Belohnungssystem kann unlogisches Verhalten auslösen³⁸³, für die wiederholte Abfolge von (Nor-)Adrenalin und Dopamin wird sogar eine hohe Gefahr der Suchtentstehung gesehen³⁸⁴. In der Konsequenz dürften auch Steuerhinterziehungen der typischen Dynamik einer zunehmenden Intensivierung von Gesetzesbrüchen folgen. Dafür spricht außerdem der aus Studien bekannte Zusammenhang, dass Menschen, die ihr Geld (spielerisch) leicht verdienen, stärker zur Steuerhinterziehung neigen, als jene, die dafür hart arbeiten müssen³⁸⁵.

³⁸⁰ Vgl. z.B.: *Ernst*, „Indirekte Methoden der steuerlichen Betriebsprüfung: klassische Verprobung, Systemprüfung, neuere Verfahren“, Schmidt-Verlag, 1986, oder *Schumann*, „Voraussetzungen und Verfahren der abgabenrechtlichen Schätzung nach § 162 Abs. 1 und 2 AO – insbesondere im Fokus der steuerlichen Außenprüfung“, Dissertation an der Johannes Gutenberg-Universität Mainz, 2009, 105 f. mit diversen Abb.

³⁸¹ Vgl. z.B.: *Damasio*, „Descartes' Irrtum. Fühlen, Denken und das menschliche Gehirn“, List/Ullstein-Verlag, 2004, 25 ff., oder *Ramachandran*, „Eine kurze Reise durch Geist und Gehirn“, Rororo-Verlag, 2. Auflage 2006, 100 f. u. 119.

³⁸² Der Nervenbotenstoff „Noradrenalin“ ist namensgebend für stressartige Reaktionen und den betreffenden Teil des (autonomen) vegetativen Nervensystems.

³⁸³ Für Funktionsstudien des Belohnungszentrums wurde Ratten per Schalter die Möglichkeit gegeben, die betreffende Zwischenhirnregion elektrisch zu stimulieren. Die Tiere haben daraufhin sämtliche (lebenserhaltenden) Handlungen (Nahrungsaufnahme, Interaktion und Sex) komplett vernachlässigt, weil sie innerhalb kürzester Zeit süchtig nach dem Schalter geworden waren, vgl.: *Olds/Milner*, „Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain“, *Journal of Comparative Physiological Psychology* 47(6), 1954, 419.

³⁸⁴ Als Alltagsbeispiel für die Suchtgefahr durch gegenseitige Verstärkung von Adrenalin, Endorphinen und Dopamin kann Extremsport angeführt werden; vgl. hierzu z.B.: news.de online v. 28.7.2009, „Faszination Extremsport: Die Sucht nach der Leere im Gehirn“, nachzulesen z.B. unter: www.news.de/gesundheit/854672263/die-sucht-nach-der-leere-im-gehirn/1.

³⁸⁵ *Tenzer*, „Warum Menschen Steuern zahlen – oder eben nicht“, *Psychologie Heute* 9/2013, 30, 32 mit Verweis auf: *Mühlbacher/Kirchler*, „Arbeitsaufwand, Anspruchsniveau und Steuerehrlichkeit“, *Zeitschrift für Arbeits- und Organisationspsychologie* 2/2008, 91.

Mehrere **Manipulationsszenarien**, die sich aus den physiologischen Abläufen oder operativen Notwendigkeiten³⁸⁶ ergeben, hinterlassen Spuren von Spitzeneingriffen und können somit durch die SRP-Verteilungsanalyse aufgedeckt werden wie insbesondere:

-  Jede Form der Einnahmenverkürzung z.B. durch Positionslöschungen oder prozentuale Minderung(en) kann zu stärkeren Spitzeneingriffen führen, wenn das Gewissen oder der Kassenbestand einen vergleichbaren Eingriff in einnahmeschwächere Tage verhindert.
-  Eine systematische Einnahmenverkürzung bspw. durch prozentuale Positionsminderungen kann dann erkennbare Spuren in den (betriebsbezogen) höheren Tageskasseneinnahmen erzeugen, wenn die Gier innerhalb der langen Phase eines Prüfungszeitraums zunimmt und der Minderungsanteil erhöht wird.
-  Bei dem typischen Steuerhinterziehungsmodell „Schwarzlöhne aus Schwarzeinnahmen“ kann es zu stärkeren Spitzeneingriffen kommen, weil die schwarz bezahlten Kosten für Aushilfskräfte an Hocheinnahmetagen oft überproportional ansteigen.

Eine permanent systematische Einnahmenverkürzung unter Zuhilfenahme digitaler Manipulationsinstrumente kann auch durch die SRP-Verteilungsanalyse nicht aufgedeckt werden. Allerdings widerspricht ein derartiges Verhalten den (kriminal-)psychologischen Erkenntnissen, weshalb *Tenzer* fragt: „Was können Psychologen den Finanzämtern raten?“³⁸⁷

4.3.4.1. Die (Punkt-)Erwartung „Logarithmische Normalverteilung“

Nach **empirischen**³⁸⁸ **Erfahrungen** findet sich die Logarithmische Normalverteilung – neben der Normalverteilung – in vielen Wissenschaftsbereichen als regelmäßig beobachtetes Verteilungsverhalten von Daten wieder.³⁸⁹ Besonders werden hierbei die Bereiche „Biologie und Wirtschaft“³⁹⁰ herausgestellt. Letzteres passt dazu, dass das Benford-Gesetz³⁹¹ ebenfalls vornehmlich in Wirtschaftsdaten wiederzufinden ist und ein inhaltlicher Zusammenhang zwischen beiden (Punkt-)Erwartungen besteht³⁹². Entsprechend berichtet *Rau*, dass

³⁸⁶ Auch Dokumentationspflichten wie das taggenaue Führen einer Kasse kann Manipulationen beeinflussen, bspw. um das Entstehen von Kassenfehlbeträgen zu vermeiden.

³⁸⁷ *Tenzer*, „Warum Menschen Steuern zahlen – oder eben nicht“, *Psychologie Heute* 9/2013, 30, 33.

³⁸⁸ Empirische Erkenntnisse stammen aus Studien, die den Untersuchungsgegenstand unter nachvollziehbaren konstanten Bedingungen beobachten, und sollen theoretische Annahmen entweder bestätigen oder verwerfen. Gesicherte empirische Erkenntnisse genießen einen höheren Stellenwert in der Wissenschaft als widerstreitende theoretische Annahmen, weil sie die Realität beschreiben.

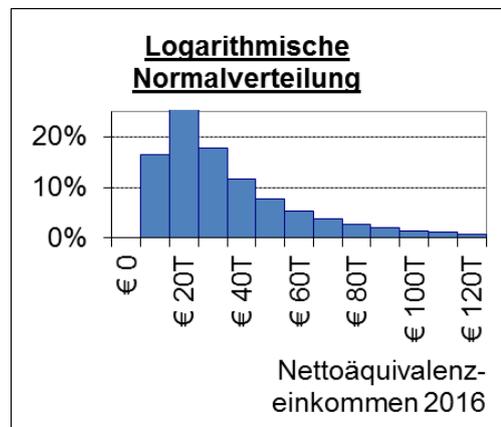
³⁸⁹ Vgl. z.B.: *Limpert/Stahel/Abbt*, „Lognormal distributions across the sciences: keys and clues“, *BioScience* Band 51, Nr. 5, 2001, 341, nachzulesen unter: https://de.wikipedia.org/wiki/Logarithmische_Normalverteilung.

³⁹⁰ *Sachs*, „Angewandte Statistik“, Springer-Verlag, 10. Auflage 2002, 173.

³⁹¹ *Benford*, „The Law of Anomalous Numbers“, *Proceedings of the American Philosophical Society* 1938, 551.

³⁹² *Petersen*, „Benford’s Law und die Betriebsprüfung – Eine Erwiderung auf Wähnerts „Interaktive Außenprüfung – Möglichkeiten der gemeinsamen Sachverhaltsaufklärung im digitalen Zeitalter“ (Stbg 2014 S. 20 ff.) aus wissenschaftshistorischer Perspektive“, *Stbg* 12/2015, 506, 510: „Ist die Kongruenz der diskreten Benford-Verteilung mit der stetigen logarithmischen Normalverteilung wohl in der Tat existent [...]“ mit Verweis auf:

„die logarithmische Normalverteilung [...] bei der Verteilung des Werts von Hauseigentum, von Industrien eines Landes nach Anzahl der Arbeitnehmer, Größe der Betriebe, Produktion, Kapital und Fabrikationskosten, bei der Verteilung von Einlagen bei Banken und Kassen, der Größe von landwirtschaftliche[n] Anwesen, von Löhnen und Gehältern in Berufsgruppen bei einzelnen Großbetrieben sowie bei vielen wirtschaftlichen Zuständen auftritt“³⁹³. So beschreibt z.B. die nebenstehende Logarithmische Normalverteilung³⁹⁴ die Einkommensverteilung 2016 in Deutschland auf Basis des Nettoäquivalenzeinkommens mit der sozialen Absicherung links und der weit auslaufenden Spreizung zu hohen Einkommen hin.



Als **mathematisch theoretische** Erklärung für das asymmetrische Streuungsverhalten der Logarithmischen Normalverteilung wird ein multiplikatives Zusammenwirken von Zufallsgrößen angeführt.³⁹⁵ M.E. dürfte jedoch der entscheidende Unterschied zur glockenförmig symmetrischen Zufallsstreuung gem. Normalverteilung in einer linksseitigen Begrenzung der möglichen Datenausprägung liegen: Als Beispiel für den technischen Bereich lassen sich unterhalb eines minimalen Kraftstoffverbrauchs keine Autos herstellen, was die untere Grenze bildet. Mit auf der X-Achse ansteigendem Kraftstoffverbrauch nehmen die Pkw-Typen bis zu einem (für den Beobachtungszeitpunkt) typischen Normverbrauch zu. In den darüber hinausgehenden Verbrauchsklassen gehen die Häufigkeiten langsamer zurück und reichen erheblich weiter bis zum Verbrauchsmaximum. Ähnliche Verteilungsdynamiken kommen in der Biologie vor, wenn bspw. Lebensmittel konserviert werden sollen. Die mikrobiologische Verunreinigung des Produkts wird durch eine (fast) erreichte Sterilität links auf der X-Achse begrenzt. Vom häufigen Standardzustand der Lebensmittel nach der Konservierung reicht die Streuung zu höheren Restverunreinigungen deutlich weiter nach rechts hinaus. Diese Form der eigentlich glockenförmigen Streuung, die aber linksseitig eingeschränkt wird, ist auch im Wirtschaftsleben typisch. Hier stammt die untere Grenze entweder aus der sozialen Absicherung z.B. eines Mindesteinkommens oder aus der Marktbereinigung, wenn

Scott/Fasli, „Benford’s law: An Empirical Investigation and a Novel Explanation“, CSM Technical Report 329, Essex 2001, 15.

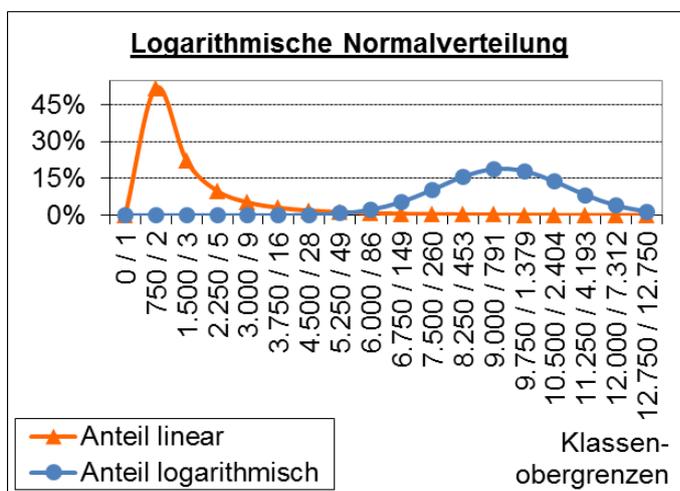
³⁹³ Rau, „Statistisch-mathematische Methoden der steuerlichen Betriebsprüfung und die Strukturanalyse als ergänzende Alternative“, Reihe: „Steuern, Wirtschaft und Recht“, Bd. 316, Eul-Verlag, 2012, 43 mit Verweis auf: Daeves/Beckel, „Großzahl-Methodik“, Verl. Chemie, 1958, 74 ff.

³⁹⁴ Die Logarithmische Normalverteilung ist auf Basis der Daten des Statistischen Bundesamtes rekonstruiert worden; die Daten können nachgelesen werden unter:
https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/EinkommenKonsumLebensbedingungen/LebensbedingungenArmutsgefahr/Tabellen/Einkommensverteilung_SILC.html.

³⁹⁵ Sachs, „Angewandte Statistik“, Springer-Verlag, 10. Auflage 2002, 173.

erfolgreiche Unternehmen geschlossen werden müssen, an absatzschwächeren Tagen nicht mehr geöffnet wird, schlecht verkäufliche Produkte aus dem Sortiment genommen werden usw. Insofern ist die regelmäßig beobachtete Häufigkeitsverteilung „Logarithmische Normalverteilung“ in Tageskasseneinnahmen mit dem Benford-Gesetz für die führenden Zahlenpositionen und der (annähernden) Gleichverteilung für endständige Zahlenstellen (ab der Drittziffer) auch theoretisch nachvollziehbar. **Wichtig ist allerdings, dass die untersuchten Daten einen einheitlichen Bestand bilden.** Die vorbereitende Werteklassifizierung der SRP-Systematik schützt davor, die mögliche Uneinheitlichkeit einer Datenmenge zu übersehen und hilft bei der Ermittlung der Ursache (s. [Pkt. 4.3.3.3.](#)).

Das Besondere an Logarithmischen Normalverteilungen gegenüber anderen asymmetrisch linkssteil-rechtsschiefen Häufigkeitsverteilungen ist, dass sie sich **unter logarithmischer Werteklassifizierung in eine symmetrische Normalverteilung verwandeln.** In der Abb.³⁹⁶ ist einmal die **asymmetrische Form einer Logarithmischen Normalverteilung** zu sehen, welche aus der Auszählung der



Anteil linear

Anteil logarithmisch

Klassenobergrenzen

Daten nach den linearen Klassen der unteren X-Achsenbeschriftung hervorgeht. Werden **dieselben** Daten entsprechend der oberen X-Achsenbeschriftung in logarithmische³⁹⁷ Intervalle aufgeteilt, ergibt sich das **glockenförmige Muster einer Normalverteilung**. Die Daten weisen eine Häufungskonzentration im unteren Wertebereich um 750 auf, wodurch **die hohe Kurvenspitze** entsteht. Da die logarithmischen Größenklassen anfangs klein sind und dann immer größer werden, erreicht unter dieser Darstellung kein Intervall vergleichbar hohe Häufigkeitsanteile. Für die vollständige Beschreibung einer Logarithmischen Normalverteilung reichen mit den zwei Kenngrößen „Mittelwert“ und „Standardabweichung“³⁹⁸ wenige Informationen aus dem dazugehörigen Datenbestand. Demnach lassen sich die Verhältnisse des Einzelfalls bei einer Prüfung von Tageskasseneinnahmen auf Spuren möglicher Spitzeneingriffe per Mustervergleich „SRP-Verteilungsanalyse“ problemlos übertragen. Die Gegenüberstellung von „Beobachtung“ und „(Punkt-)Erwartung“ unter logarithmischer Werteklassifizierung i.R. der SRP (s. [Pkt. 4.3.4.2.](#)) hat zwei Vorteile:

³⁹⁶ Aus Übersichtlichkeitsgründen werden beide klassifizierten Verteilungen als Liniengrafen dargestellt.

³⁹⁷ Die Klassenobergrenzen entsprechen ca. $10^{0,24}$ – $10^{0,48}$ – $10^{0,72}$ – $10^{0,96}$ etc.

³⁹⁸ Die Standardabweichung beschreibt die Streuungsintensität einer Datenmenge als vorzeichenneutrale durchschnittliche Abweichung jedes einzelnen Datensatzes zum Mittelwert.

1. Das visuelle Ergebnis der Verteilungsanalyse kann aufgrund der Symmetrie der (Punkt-) Erwartung leichter beurteilt werden.
2. Infolge der permanent größer werdenden (logarithmischen) Häufigkeitsklassen kann die Übereinstimmungsprüfung mittels χ^2 -Test (Chi-Quadrat-Test) bis zu deutlich höheren Werten mit einem größeren Manipulationsrisiko (s. [Pkt. 4.3.4.](#)) ausgewertet werden, ohne die Mindestdatenempfehlung (s. [Pkt. 4.3.1.3.](#)) zu verletzen.

Die Übereinstimmungsprüfung des beobachteten Häufigkeitsmusters aus der Werteklassifizierung der untersuchten Tageskasseneinnahmen mit der Logarithmischen Normalverteilung erfolgt durch den χ^2 -Test (Chi-Quadrat-Test). Dadurch beinhaltet das Übereinstimmungsergebnis die Datenanzahl (s. [Pkt. 4.3.1.3.](#)). Die übliche Empfehlung zur Mindestdatenmenge „ ≥ 5 erwartete Elemente pro Klasse“ wird von der SRP-Vorlage Stochastik automatisch in Abhängigkeit von der Klassenfestlegung beachtet, indem nur der Hauptverteilungsbereich (stochastisch) ausgewertet wird³⁹⁹. Auf dem Tabellenblatt „Verteilungsanalyse“ sind die einbezogenen und die ausgeschlossenen Intervalle leicht zu unterscheiden (s. *das Beispiel im Abschnitt „Chi-Quadrat-Test“ unter [Pkt. 4.3.4.2.](#)*). Die zwei Auffälligkeitsschwellen „99 %“ und „99,95 %“ werden formatseitig hervorgehoben (s. [Pkt. 4.3.2.1.](#)).

4.3.4.2. Anwendungshinweise zur SRP-Verteilungsanalyse

Folgende Hinweise zur Durchführung einer SRP-Verteilungsanalyse mit der SRP-Vorlage Stochastik sind wichtig:

- ➔ Alle auf dem Tabellenblatt „Daten“ vorgenommenen **generellen Dateneinschränkungen** (s. *den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)*) wirken auch auf die Verteilungsanalyse der SRP-Vorlage Stochastik durch. Ein entsprechender Hinweis erfolgt auf den Auswertungsblättern „Verteilungsanalyse“ und „Verteilungsdiagramm“.
- ➔ Für die empfohlene automatische Klassenfestlegung der Verteilungsanalyse werden ebenfalls die erforderlichen statistischen Kenngrößen auf dem Tabellenblatt „Daten“ unterhalb derjenigen für die linearen Werteklassifizierungen ausgegeben (s. *den nächsten Aufzählungspkt. „Klassenunterteilung mit der Wurzelmethode“*). Enthalten die zu untersuchenden Daten **negative oder Nullwerte** wird durch rote Feldfarbe zur „Anzahl“ auf die Abweichung zwischen Schichtung und Verteilungsanalyse hingewiesen. Da kein Logarithmus zu negativen und Nullwerten existiert, können diese Datensätze in die Verteilungsanalyse nicht mit einbezogen werden. *Als speziell für die Untersuchung von Ta-*

³⁹⁹ Kurvenverteilungen, die sich (asymptotisch) der X-Achse annähern, wie die Normal- und die Logarithmische Normalverteilung fallen zwangsweise zu den Seiten hin irgendwann unter die Klassengrenze „ ≥ 5 “. Deshalb wird der χ^2 -Test (Chi-Quadrat-Test) als Übereinstimmungsprüfung für die Logarithmische Normalverteilung i.R. der Verteilungsanalyse auf die Klassen mit einer Erwartung „ ≥ 5 “ beschränkt.

geskasseneinnahmen auf „Spitzeneingriffe“ (s. [Pkt. 4.3.4.](#)) *ausgestalteter Prüfungsschritt bedeutet das Vorliegen von negativen und Nullwerten ein eigenes Prüffeld.*

- ➔ Die **Klassenunterteilung mit der Wurzelmethode**⁴⁰⁰ folgt demselben **Ablauf** wie bei den Werteklassifizierungen (s. *den Aufzählungspkt. „Klassenunterteilung mit Wurzelmethode“ unter [Pkt. 4.3.3.1.](#)*). Um jedoch die Vorteile einer logarithmischen⁴⁰¹ Klassifizierung (s. [Pkt. 4.3.4.1.](#))

nutzen zu können, werden die statistischen Kenngrößen der zu untersuchenden Daten als Logarithmen zur Basis „10“⁴⁰² wiedergegeben. Nach jeder Dateneinschränkung auf dem Tabellenblatt „Daten“ (s. *die*

statistische Kenngrößen	
logarithmische Kenngrößen (Verteilungsanalyse)	
Anzahl:	705
Mittelwert:	2,896
Standardabweichung:	0,183
Auffälligkeitsmaß (StdNV für +1):	3,062
Minimum:	2,145
Hilfsminimum:	2,315
Maximum:	3,509
Hilfsmaximum:	3,505
Vorschlag Intervallanzahl:	27
Vorschlag Schichtungsintervall:	0,050
Vorschlag erste Schichtobergrenze:	2,320
Vorschlag zweite Schichtobergrenze:	2,370

zweite Abb. zu „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) wird die Klassifizierung der Verteilungsanalyse automatisch umgehend angepasst. Neben der empfohlenen Klassenbildung können die zwei ersten Obergrenzen (und damit die Intervallgröße) manuell abgeändert werden (s. *den übernächsten Aufzählungspkt. zur „freien Klassenbildung“*).

- ➔ Die SRP-Vorlage Stochastik versucht für zielführendere Werteklassifizierungen extreme Einzelwerte ohne engeren Verteilungsbezug zu den untersuchten Daten (Ausreißer) selbständig zu erkennen und verwendet in dem Fall die Visualisierungsgrenzen „**Hilfsminimum**“ und „**Hilfsmaximum**“. Dazu werden die Erfahrungen nach *Grubbs*⁴⁰³ abgewandelt herangezogen: Datensätze, welche mehr als den fünffachen Interquartilsabstand⁴⁰⁴ nach oben oder unten vom Median⁴⁰⁵ abweichen, gelten als „Ausreißer“. Erkannte Ausreißer werden durch die Hilfswerte als Visualisierungsgrenzen der Vertei-

⁴⁰⁰ Die Wurzelmethode berechnet die Klassenanzahl zur Klassifizierung (Schichtung) nach folgender Formel: $\sqrt{(\text{Anzahl Datensätze})}$. Das Ergebnis wird in der SRP-Vorlage Stochastik zur besseren Übersichtlichkeit auf max. 30 Intervalle begrenzt.

⁴⁰¹ Eine logarithmische Skalierung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie bspw. mit „10⁰ – 10¹ – 10² – 10³“ (entsprechend „1 – 10 – 100 – 1.000“).

⁴⁰² Eine Tageskasseneinnahme von € 1.000 entspricht somit der Zahl „3“, weil gilt: $\log_{10} 1.000 = 3$, also $10^3 = 1.000$.

⁴⁰³ *Grubbs*, „Engineering Statistics Handbook“, Kapitel 7.1.6. „What are outliers in the data?“, nachzulesen unter: <https://www.itl.nist.gov/div898/handbook/prc/section1/prc16.htm>.

⁴⁰⁴ Der Interquartilsabstand einer Datenmenge beschreibt den Wertumfang zwischen dem 25 %- und dem 75 %-Quantil (1. und 3. Quartil). Dazu wird die Differenz zwischen dem Wertinhalt des – aufsteigend sortiert – 75 %- und 25 %-größten Rangs ermittelt.

⁴⁰⁵ Der Median ist der mittlere, also der 50 %-Rang einer Datenmenge.

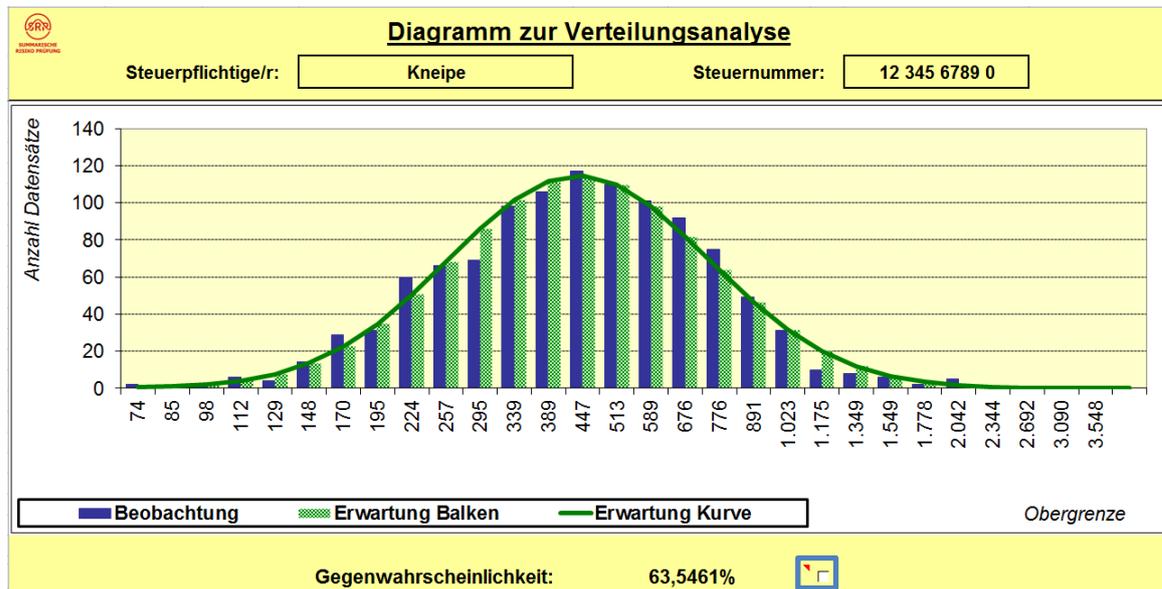
lungsanalyse ersetzt. Dies wird aus dem logarithmischen Teil der „statistischen Kenngrößen“ auf dem Tabellenblatt „Daten“ ersichtlich, indem die Werte „Minimum“ und „Maximum“ hellgrau unterdrückt sind. Die betreffenden Ausreißer-Datensätze werden dadurch **nicht** aus den Analysen ausgeschlossen. Ist dies gewünscht, weil es sich z.B. um eine nicht mit normalen Tageskasseneinnahmen vergleichbare Tagesveranstaltung handelt, kann der Datensatz per Kreuz in der Kästchenspalte auf dem Tabellenblatt „Daten“ abgewählt werden (s.a. den Abschnitt „Hilfsminimum und Hilfsmaximum“ unter [Pkt. 4.3.3.1.](#)). In dem Fall erscheint auf jedem Auswertungsblatt automatisch der Hinweis „Untersuchte Daten: Einzeldatenabwahl“.

- ➔ Die Datenklassifizierung der Verteilungsanalyse kann auf dem Tabellenblatt „Verteilungsanalyse“ auch mit **freier Klassenbildung** durchgeführt werden. Dazu lassen sich die ersten zwei Schichtobergrenzen in den beiden freien Feldern oberhalb einzeln oder insgesamt abändern. Die **eingetragenen Exponenten zur Basis „10“** werden in der Spalte daneben in den Wert übersetzt. Das durch die Eingabe(n) entstehende Intervall wird in der SRP-Vorlage Stochastik automatisch bis zum höchsten Wert der untersuchten Daten bzw. zur maximalen Klassenanzahl „30“ fortgesetzt.

Schichtobergrenze	
lg	absolut
→ 1,950	
1,870	74
1,950	89
2,030	107
2,110	129
2,190	155

- ➔ Falls infolge der (automatischen oder einer manuell abgeänderten) Klassenfestlegung **nicht alle hohen Spitzenwerte in der Auswertung berücksichtigt** werden können, verfärbt sich das Summenfeld unter der Häufigkeitsverteilung auf dem Tabellenblatt „Verteilungsanalyse“ rot.
- ➔ Die Funktion „**Absolutwerte**“ zu den Werteklassifizierungen auf dem Tabellenblatt „Schichtung“ (s. den Aufzählungspkt. „*vorzeichenneutrale Beträge*“ unter [Pkt. 4.3.3.1.](#)), mit welcher Korrekturbuchungen bzw. -aufzeichnungen hervorgehoben werden, hat auf die Herleitung der betriebsspezifischen Logarithmischen Normalverteilung i.R. der Verteilungsanalyse keine Auswirkung. Deren Ausrichtung „Untersuchung von Tageskasseneinnahmen auf ‚Spitzeneingriffe‘“ (s. [Pkt. 4.3.4.](#)) stünde dazu im Widerspruch.
- ➔ Die Festlegung eines „**Startwerts**“ für die Werteklassifizierungen auf dem Tabellenblatt „Schichtung“ (s. den Aufzählungspkt. „*Startwert*“ unter [Pkt. 4.3.3.1.](#)) hat auf die Verteilungsanalyse ebenfalls keine Auswirkung, weil deren Ausrichtung „Untersuchung von Tageskasseneinnahmen auf ‚Spitzeneingriffe‘“ (s. [Pkt. 4.3.4.](#)) auch hierzu im Widerspruch stünde.
- ➔ Zur besseren Übersichtlichkeit wird das **visuelle Ergebnis** der Gegenüberstellung von ausgezählten Tageskasseneinnahmen und der einzelfallbezogen abgeleiteten Loga-

rithmischen Normalverteilung in der SRP-Vorlage Stochastik vom Übereinstimmungstest (s. *nächster Aufzählungspkt.*) getrennt auf dem eigenen Tabellenblatt „Verteilungsdiagramm“ ausgegeben:



Durch die logarithmische⁴⁰⁶ Werteklassifizierung auf der X-Achse wird der optische Abgleich zur symmetrischen (Punkt-)Erwartung erleichtert; dass es sich um eine logarithmische Steigerung der Intervallgröße handelt, kann auf dem Tabellenblatt „Verteilungsanalyse“ nachvollzogen werden. Ebenso dient die doppelte Darstellung der erwarteten Logarithmischen Normalverteilung als Balken- und Liniengraf⁴⁰⁷ dem verbesserten optischen Erkennen von Übereinstimmung und Abweichungen. Vom Übereinstimmungstest des Tabellenblatts „Verteilungsanalyse“ wird nur das Übereinstimmungsergebnis des χ^2 -Tests (Chi-Quadrat-Test) übertragen und lässt sich über ein **Anhakfeld** bei Bedarf ausblenden.

- ➔ Die qualifizierte **Übereinstimmungsprüfung** der Tageskasseneinnahmen mit der (einzelfallspezifischen) Logarithmischen Normalverteilung unter Berücksichtigung der untersuchten Datenmenge erfolgt in der SRP bewusst⁴⁰⁸ mit dem χ^2 -Test (Chi-Quadrat-Test). Das komplette Zahlenwerk dazu findet sich auf dem Tabellenblatt „Verteilungsanalyse“:

⁴⁰⁶ Eine logarithmische Skalierung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie bspw. mit „10⁰ – 10¹ – 10² – 10³“ (entsprechend „1 – 10 – 100 – 1.000“).

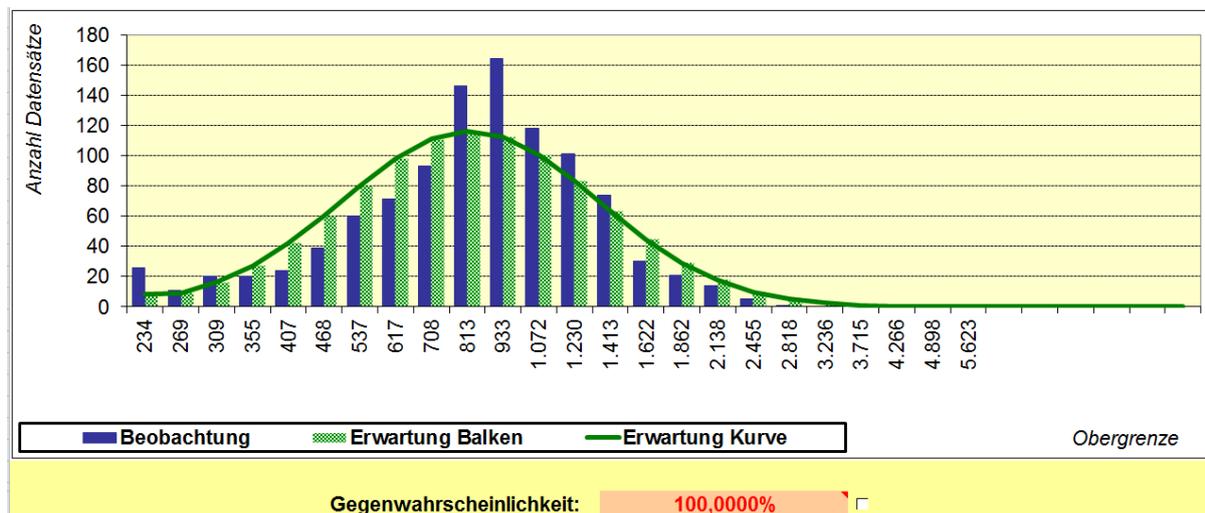
⁴⁰⁷ Die Darstellung der (diskreten) logarithmischen Werteklassifizierung als kontinuierlicher Liniengraf ist streng genommen nicht richtig, trägt aber zu einer korrekteren optischen Beurteilung bei.

⁴⁰⁸ Symmetrietests würden die Auslaufzonen mit (sehr) wenigen Elementen im Übereinstimmungsergebnis gleich stark werten wie den *einzelfalltypischen* Häufungsbereich, obwohl die Auswirkungen einzelner Datensätze an der Grenze von zwei Intervallen hier besonders groß sind. I.R. des χ^2 -Tests (Chi-Quadrat-Test) lässt sich dieses Problem durch die Einhaltung der Mindestdatenempfehlung (s. *Pkt. 4.3.1.3.*) einfach lösen.

angegeben. Da im Gegensatz zur Einerstellenuntersuchung und Zweitiffernanalyse die Klassenanzahl⁴¹³ bei der Verteilungsanalyse variiert, verschiebt sich auch der Zusammenhang zwischen χ^2 -Gesamtwerten und (Nicht-)Übereinstimmungswahrscheinlichkeiten. Der gesondert angegebene Abweisungswert benennt den χ^2 -Gesamtwert, welcher bei der Ziffernanalyse (mit zehn Ziffernmöglichkeiten) „30“ entspricht. Dieser Grenzwert wird als Klassenabweisungswert auf die Anzahl ausgewerteter Intervalle gleichmäßig verteilt, um in der letzten Spalte die besonders kritischen Bereiche der untersuchten Tageskasseneinnahmen zu markieren: Klassen mit einem zu hohen χ^2 -Einzelwert, welche betriebsbezogen zu den hohen⁴¹⁴ Einnahmen zählen und unterbesetzt sind wie in Abb. oberhalb „> € 1.023 bis € 1.175“.

4.3.4.3. Ergebnisverdeutlichung bei auffälligen Verteilungsanalysen

Um die wahrscheinlichste und kriminalpsychologisch typische Manipulationsspur „Spitzeneingriffe“ (s. [Pkt. 4.3.4.](#)) in Tageskasseneinnahmen eines entsprechenden Falls deutlich hervorzuheben, verfügt die SRP-Vorlage Stochastik über eine **gezielte Anpassung der betriebsspezifischen Logarithmischen Normalverteilung**.



Im abgebildeten Beispielsfall eines gastronomischen Betriebs weicht die aus den Kennzahlen „Mittelwert“ und „Standardabweichung“⁴¹⁵ der Tageskasseneinnahmen rekonstruierte (betriebsspezifische) Logarithmische Normalverteilung von dem beobachteten Häufigkeitsmuster mathematisch eindeutig (signifikant) ab. Den zu großen Häufungen in der untersten

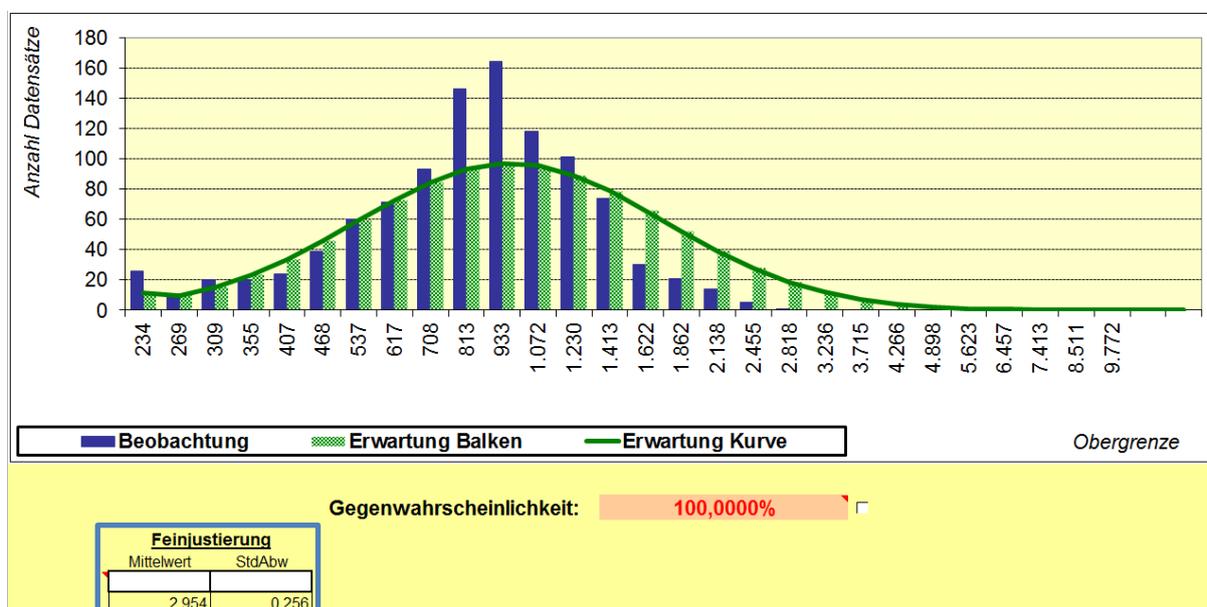
⁴¹³ Die Klassenanzahl einer Verteilungsgegenüberstellung, die mit dem χ^2 -Test (Chi-Quadrat-Test) auf Übereinstimmung geprüft werden soll, findet sich in der erforderlichen Information „Freiheitsgrade“ wieder: i.d.R. als „Klassenanzahl – 1“.

⁴¹⁴ Die Schwelle ergibt sich mit der Klassenoberschicht, die größer als „Mittelwert + Standardabweichung“ aus den untersuchten Tageskasseneinnahmen ist.

⁴¹⁵ Die Standardabweichung beschreibt die Streuungsintensität einer Datenmenge als vorzeichenneutrale durchschnittliche Abweichung jedes einzelnen Datensatzes zum Mittelwert.

Werteklasse und in den Intervallen „> € 708 bis € 1.413“ stehen linksseitig und rechtsseitig davon Unterbesetzungen gegenüber.

Falls es sich bei den festgestellten Unterschieden um stärkere Eingriffe in die betriebsbezogen höheren Einnahmen handelt, ist auch der Mittelwert hiervon betroffen⁴¹⁶. Somit kann diese abgeleitete Logarithmische Normalverteilung nicht die ursprünglichen Absatzverhältnisse wiedergeben. Nach kriminalistischen und Prüfungserfahrungen tragen die Tageskasseneinnahmen in der Anstiegszone „> € 355 bis € 708“ ein deutlich geringeres Manipulationsrisiko als die höheren Werte. Deshalb bietet die SRP-Vorlage Stochastik die Möglichkeit an, per Auswahl der Funktion „angepasst“ (s. die Abb. zum Aufzählungspkt. „Übereinstimmungsprüfung“ unter [Pkt. 4.3.4.2.](#)) aus einer aussagekräftigen Datenanzahl dieses Steigungsbereichs⁴¹⁷ die betriebsspezifische Logarithmische Normalverteilung abzuleiten:



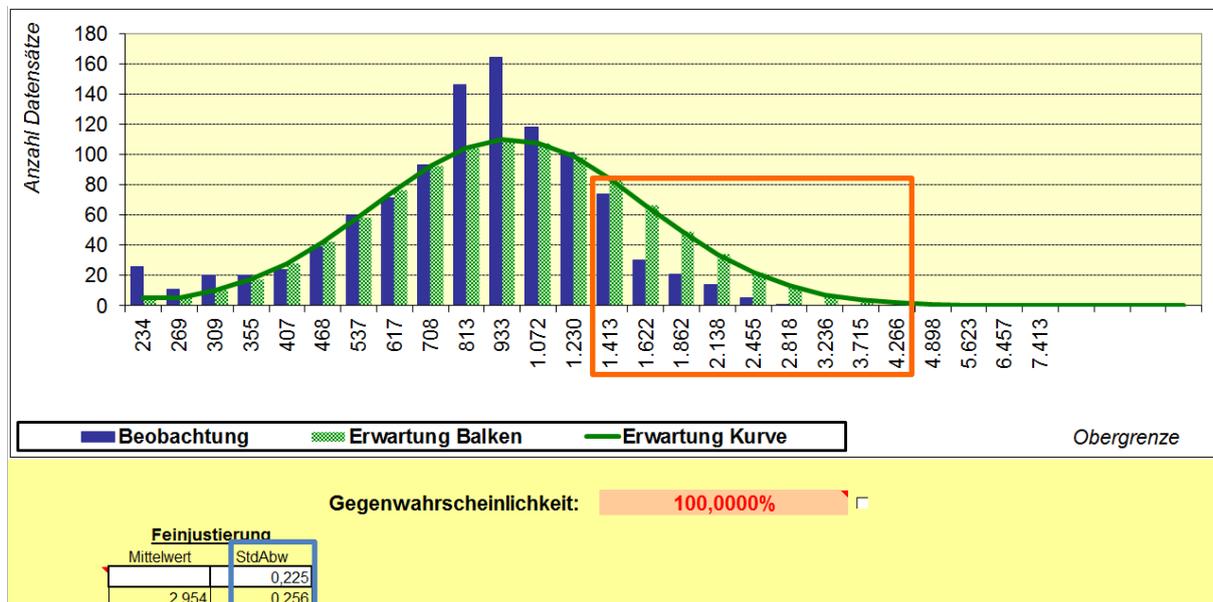
In der angepassten Gegenüberstellung ist gut zu erkennen, dass die – weiterhin betriebsintern rekonstruierte – Logarithmische Normalverteilung mit der Häufigkeitszunahme im Bereich „> € 355 bis € 708“ die weniger hinterziehungsanfälligen Verhältnisse des Einzelfalls bereits weitgehend trifft. Um die Nachfragesituation des geprüften Betriebs aus den Tageskasseneinnahmen so genau wie möglich wiederherzustellen, können die beiden **Kenngößen „Mittelwert“ und „Standardabweichung“** **manuell feinjustiert** werden:

⁴¹⁶ Da bei einer Einnahmenverkürzung die Summe der Einnahmen reduziert wird, sich die Anzahl der Tage aber i.d.R. nicht (in gleichem Maß) durch komplettes Streichen verringert, wird der Mittelwert als $\frac{\text{Summe}}{\text{Anzahl}}$ durch die Manipulation ebenfalls verkleinert.

⁴¹⁷ Mit der Steigungszone zu den Werten „-0,75“ und „-0,5“ aus der Standardnormalverteilung werden ca. das 23 %- und 31 %-Quantil gewählt. So verzerren weder mögliche Ausreißer im untersten Wertebereich noch starke Konzentrationen in den vertrauten Haupthäufungsklassen den hochgerechneten Mittelwert.

- ➔ Eine Erhöhung des Mittelwerts verschiebt die erwartete Logarithmische Normalverteilung auf der X-Achse nach rechts. Demgegenüber bedeutet eine Minderung eine Verschiebung der gesamten Kurvenstruktur nach links.
- ➔ Durch eine Vergrößerung der Standardabweichung wird die Logarithmische Normalverteilung auseinandergezogen, so dass die Häufigkeiten in der Mitte ab- und an den Rändern zunehmen. Entsprechend bewirkt eine Absenkung der mittleren Abweichung vom Mittelwert eine Stauchung der Kurvenstruktur mit stärkerer Häufigkeitskonzentration in der Mitte.

Im vorliegenden Beispielfall ist eine geringfügige **Reduzierung der Standardabweichung** angezeigt, um die etwas stärkere Steigung der Häufigkeiten in der Zone „> € 355 bis € 708“ besser wiederzugeben:



Nach der betriebsnahen Entwicklung der (Punkt-)Erwartung „Logarithmische Normalverteilung“ an einer größeren Anzahl Tageskasseneinnahmen aus dem weniger manipulationsanfälligen Wertebereich bleibt es lt. χ^2 -Test (Chi-Quadrat-Test) (für die auswertbaren Klassen) bei einer mathematisch eindeutigen (signifikanten) Nichtübereinstimmung.

Nun lässt sich dem Diagramm das manipulationstypische (s. [Pkt. 4.3.4.](#)) Szenario entnehmen: Falls die unteren Besteuerungsdaten richtig sind, müsste die Absatzsituation des Unternehmens gem. den Erkenntnissen zur Verteilung von Wirtschaftsdaten (s. [Pkt. 4.3.4.1.](#)) in den mittleren und hohen Einnahmeregionen der grünen Kurve folgen. Tatsächlich ist eine unerwartet intensive Häufung besonders in dem Standardbereich „> € 708 bis € 933“ auf Kosten aller hohen Intervalle zu beobachten. Übertragen auf den Verdacht einer Einnahmenverkürzung resultiert eine derartige Abweichung von der Logarithmischen Normalverteilung dadurch, dass regelmäßig (stärker) in die betriebsbezogen hohen Tageskasseneinnah-

men eingegriffen worden ist und dabei häufig Werte im vertrauten Standardbereich entstanden sind.

Im Gegensatz zu Ziffernanalysen, die nur einzelne Zahlenpositionen in den untersuchten Daten betrachten, lässt sich aus einer auffälligen Verteilungsanalyse mit manipulationspsychologisch typischer Musterabweichung auch eine Quantifizierung ableiten (s. [Pkt. 4.3.4.6.](#)).

Alle auf dem Tabellenblatt „Daten“ vorgenommenen **generellen Dateneinschränkungen** (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) wirken auch auf die angepasste Verteilungsanalyse der SRP-Vorlage Stochastik durch. Ein entsprechender Hinweis erfolgt auf den Auswertungsblättern „Referenzsumme“ und „Referenzdiagramm“ (s. [Pkt. 4.3.4.6.](#)).

4.3.4.4. Nicht-manipulative Ursachen für auffällige Verteilungsanalysen

Folgende **einzelbetriebliche Besonderheiten** bzw. andersartige Eingriffe können ein Häufigkeitsmuster verursachen, welches in der Verteilungsanalyse mathematisch eindeutig (signifikant) auffällt und dabei evtl. einem manipulationspsychologisch typischen Spitzeneingriff (s. [Pkt. 4.3.4.](#)) ähnelt:

1. Beim **freien Erfinden von Steuerungsdaten** fehlt der wirtschaftliche Entstehungszufall, wie ihn bspw. die Nachfrage entstehen lässt. Insofern werden frei erfundene Tageskasseneinnahmen häufig eine chaotische Häufigkeitsverteilung hervorbringen, in der keine Gleichmäßigkeit erkennbar ist. Die doppelte Gegenüberstellung der betriebsspezifischen (Punkt-)Erwartung „Logarithmische Normalverteilung“ als Balken- und Liniengraf (s. den Aufzählungspkt. „visuelles Ergebnis“ unter [Pkt. 4.3.4.2.](#)) soll bei der Unterscheidung helfen.
2. Falls die Absatzsituation eines Betriebs zeitweise so stark vom Standard abweicht, dass es sich quasi um ein anderes Unternehmen handelt, kann eine **zwei- oder mehrgipflige Häufigkeitsverteilung** (s. die ersten beiden Abb. unter [Pkt. 4.3.3.3.](#)) resultieren. Dieser Effekt wird aus den verbildlichten Werteklassifizierungen ersichtlich: Eine Häufigkeitskonzentration stellt das übliche Spektrum an Tageskasseneinnahmen dar. Entweder davor oder in deren Auslauf erscheint ein zweiter Verteilungsanstieg mit Häufigkeitsmaximum und -abfall. *Als möglicher Anwendungsfehler der SRP-Verteilungsanalyse wird eine derartige Zweifachverteilung i.d.R. vorkommen, wenn bei Betrieben mit Verzehr an Ort und Stelle sowie Außerhausgeschäften beide Tagessummen gemeinsam in einer Analyse verarbeitet werden. Weil die beiden Absatzanteile in den meisten Fällen nicht dasselbe Standardniveau haben, ergeben sich zwei Häufigkeitsverteilungen, die sich überlagern.*

3. Der gleichmäßig langsame Auslauf in den Häufigkeiten von Tageskasseneinnahmen zu den betriebsbezogen höchsten Werten hin kann durch eine **Kapazitätsgrenze** gestört sein. Sollte es z.B. zum Geschäftsmodell gehören, eine besonders hohe Nachfrage nicht zu bedienen, wird die – unter logarithmischer⁴¹⁸ Skalierung – glockenförmige Häufigkeitsverteilung innerhalb eines bestimmten Intervalls abbrechen. Alle Tage mit noch besseren Einnahmen, die ansonsten im Auslauf verteilt wären, häufen sich dann in der begrenzenden Klasse. *Eine Sonderform an Kapazitätsgrenze ist zudem durch Abhängigkeiten denkbar, wenn z.B. Apotheken eng mit Ärzten zusammenarbeiten und diese an die Verschreibungsbeschränkungen gebunden sind.*

4.3.4.5. Abklärung von auffälligen Verteilungsanalysen

Bereits die **Beschränkung auf ein praxistypisches Manipulationsszenario** bedeutet gegenüber einer vollständigen⁴¹⁹ Interpretation des klassifizierten Häufigkeitsmusters von Tageskasseneinnahmen eine deutliche reduzierte Gefahr von fehlerhaften Anwendungen und Schlussfolgerungen. Verstärkte Spitzeneingriffe in die betriebsbezogen höheren Einnahmewerte (s. [Pkt. 4.3.4.](#)) sind nicht nur die manipulationspsychologisch wahrscheinlichste Spur (digitaler) Einnahmenverkürzungen, sondern können auch am einfachsten und eindeutigsten identifiziert werden.

Die Abgrenzung einer auffälligen Verteilungsanalyse gegen die ersten beiden *unter [Pkt. 4.3.4.4.](#)* genannten Alternativursachen wird durch das Prüfungsschema zur Vermeidung von Fehldeutungen (s. *den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)*) gewährleistet:

1. Mit Hilfe der stets vorgeschalteten **redundanten**⁴²⁰ **und variablen SRP-Ziffernanalyse** (s. [Pkt. 4.3.2.](#)) kann ein möglicher Verdacht auf frei erfundene Tageskasseneinnahmen weitgehend ausgeräumt – wenn auch nicht mathematisch eindeutig (signifikant) ausgeschlossen – werden.
2. Durch die **vorbereitende Werteklassifizierung zur Feststellung von Zwei- und Mehrfachverteilungen** (s. [Pkt. 4.3.3.3.](#)) wird der praxisrelevanteste Störfaktor eines externen Vergleichs von Tageskasseneinnahmen mit der (betriebsspezifischen)

⁴¹⁸ Eine logarithmische Skalierung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie bspw. mit „10⁰ – 10¹ – 10² – 10³“ (entsprechend „1 – 10 – 100 – 1.000“).

⁴¹⁹ Vgl. z.B.: Ernst, „Indirekte Methoden der steuerlichen Betriebsprüfung: klassische Verprobung, Systemprüfung, neuere Verfahren“, Schmidt-Verlag, 1986, oder Schumann, „Voraussetzungen und Verfahren der abgabenrechtlichen Schätzung nach § 162 Abs. 1 und 2 AO – insbesondere im Fokus der steuerlichen Außenprüfung“, Dissertation an der Johannes Gutenberg-Universität Mainz, 2009, 105 f. mit diversen Abb.

⁴²⁰ Es werden mit der Einerstelle und Zweitziffer zwei unabhängige (s.a. *den Ausschluss von Werten unter € 100 bei der Einerstellenuntersuchung unter Pkt. 4.3.2.2.1.*) Zahlenpositionen gleichzeitig untersucht.

Logarithmischen Normalverteilung im Vorfeld abgeklärt. Falls dabei ein gleichmäßiges, aber zwei- oder mehrgipfliges Häufigkeitsmuster beobachtet wird, stellt die SRP-Vorlage Stochastik zielführende Funktionen zur Verfügung (s. *das Beispiel unter [Pkt. 4.3.3.3.](#)*), um die uneinheitlichen Teildatenbestände zu identifizieren, welche ursächlich für die Mischverteilung sind. Anschließend können die unvereinbaren Teile der Tageskasseneinnahmen jeweils in einer eigenen SRP-Vorlage Stochastik untersucht werden.

Der mögliche Störeinfluss „Kapazitätsgrenze“ (s. [Pkt. 4.3.4.4.](#)) für eine Übereinstimmung von Tageskasseneinnahmen mit der (betriebsspezifischen) Logarithmischen Normalverteilung ist nach ständigen Prüfungserfahrungen eher selten und lässt sich anhand der **grafischen Auswertung zur Verteilungsanalyse** auf dem Tabellenblatt „Verteilungsdiagramm“ gut nachvollziehen: Bricht das beobachtete Häufigkeitsmuster in der höchsten Werteklasse konsequent ab, ist eine Kapazitätsgrenze als Ursache nicht auszuschließen. I.d.R. sind aber vereinzelte Tage in Intervallen mit noch höheren Einnahmen zu finden wie in der letzten Abb. *unter [Pkt. 4.3.4.3.](#)* Für diese problemlos herausfilterbaren Tage müssen abweichende Bedingungen jenseits der Kapazitätsgrenze wie bspw. erweiterte Öffnungszeiten, ein Veranstaltungs-Pauschalpreis oder die ordnungswidrige Erfassung von zwei Tagen in einer Summe vorliegen. Sollte ein bestimmter Wochentag mit andersartiger Geschäftsgrundlage als mögliche Ursache in Betracht kommen, kann er auf dem Tabellenblatt „Daten“ einfach ausgeschlossen werden (s. *den Aufzählungspkt. „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)*). Danach muss der Überhang zur Kapazitätsgrenze wegfallen. Insofern kann auch dieser nicht-manipulative Sonderfaktor eines Einzelbetriebs ohne größeren Aufwand konkret abgeklärt werden.

Die Beweiskraft von mathematisch eindeutig (signifikant) auffälligen Verteilungsanalysen wird wegen der vergleichbaren Voraussetzungen aller manipulationsorientierten Musteranalysen in einem gemeinsamen Kapitel *unter [Pkt. 4.3.5.](#)* behandelt. Die derzeitige Beschränkung für Ziffernanalysen auf den Status eines „Zusatzindizes“⁴²¹ lässt sich durch verbreitete fehlerhafte Methodenbeschreibungen⁴²² und die resultierenden Zweifel an der korrekten Anwendung⁴²³ erklären.

⁴²¹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. *Trossen*; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. *Zimmermann* [Fahrtbuch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. *Kühnen*; *Sosna*, St-Bp. 2000, 41 [Teil I], 68 [Teil II]; *Meyer*, DStR 2005, 2114; *Watrin/Stuffert*, DB 2006, 1748; *Huber*, StBp. 2009, 121 [124 ff.]“.

⁴²² Vgl. z.B.: *Bisle*, „Chi-Quadrat-Test‘ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PStR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-

4.3.4.6. Quantifizierung von auffälligen Verteilungsanalysen

Im Gegensatz zu Ziffernanalysen, bei denen einzelne Zahlenpositionen auf Manipulationspuren untersucht werden, betrachtet die SRP-Verteilungsanalyse Zahlen insgesamt. Mit der betriebspezifischen Herleitung einer Logarithmischen Normalverteilung (s. [Pkt. 4.3.4.1.](#)) als (Punkt-)Erwartung (s. [Pkt. 4.3.1.2.](#)) zu den untersuchten Tageskasseneinnahmen lässt sich aus der Verteilungsanalyse auch eine **Gesamtsumme hochrechnen**. Als betriebsnah ermittelter erwarteter Absatz ist dieser Wert *in geeigneten Fällen* (s. [Pkt. 4.3.4.3.](#)) vergleichbar mit dem „Most Likely Error (MLE)“ einer mathematisch-statistischen Stichprobenprüfung (s. den *Aufzählungspkt. „statistische Stichprobenverfahren“ im Abschnitt „wichtige bestehende Erkenntnisse zur Datenanalyse“ unter [Pkt. 2.](#)*): Zwar ist die Wahrscheinlichkeit, dass der Fehler in der Grundgesamtheit genau dem MLE entspricht extrem gering, dennoch ist sie für kein anderes Einzelergebnis höher. Deshalb entspricht der MLE einer vorgenommenen Stichprobenprüfung dem wahrscheinlichsten Gesamtfehler.

In der SRP-Vorlage Stochastik kann auf dem Tabellenblatt „Referenzsumme“ die (annäherungsweise) Berechnung des Sollwerts zu den Tageskasseneinnahmen **per Haken aktiviert** werden; *diese Funktion erfordert eine bewusste Aktivierung, weil sie nur dann anzuwenden ist, wenn das typische Manipulationsbild „Spitzeneingriff“ (s. [Pkt. 4.3.4.3.](#)) erkennbar ist und andere – insbesondere nicht-manipulative – Ursachen soweit möglich ausgeschlossen wurden* (s. [Pkt. 4.3.4.5.](#)):

Einnahmensumme der angepassten Referenzverteilung					
<input checked="" type="checkbox"/>	Steuerpflichtige/r:	Restaurant		Steuernummer:	12 345 6789 0
Schichtobergrenze	Referenzhäufigkeit (gerundet)			Beobachtung	Betrag (exakte Berechnung)
	absolut	kumuliert	schichtweise		
	150	1	1,2	7	44,97
	240	13	11,6	21	1.248,82
	330	46	33,1	36	6.367,03
	420	102	55,8	46	16.669,62
	510	174	72,6	64	29.859,54
				⋮	

Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Erkenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.

⁴²³ Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

6.090	1.036	0,1	0	338,24
6.180	1.036	0,0	0	311,42
6.270	1.037	0,0	0	286,90
6.360	1.037	0,0	0	264,46
6.450	1.037	0,0	0	243,91
6.540	1.037	0,0	0	225,09
6.630	1.037	0,0	0	207,83
6.720	1.037	0,0	0	192,00
6.810	1.037	0,0	0	177,48
optionales Maximum: <input type="text"/>				
Bemerkung(en):		1.037	1.037	1.149.099,29
		erklärt / vorangemeldet:		882.061,93
		Abweichung:		267.037,36
<input type="text" value="Untersuchte Daten: Einzeldatenabwahl"/>				

Auf der Basis des gemittelten Klassenwerts und der ebenfalls gemittelt⁴²⁴ erwarteten Klassenhäufigkeit wird nachvollziehbar (annäherungsweise) der Gesamtwertinhalt der betriebspezifisch rekonstruierten Logarithmischen Normalverteilung errechnet. Zudem wird die Differenz zu den untersuchten Tageskasseneinnahmen ausgegeben.

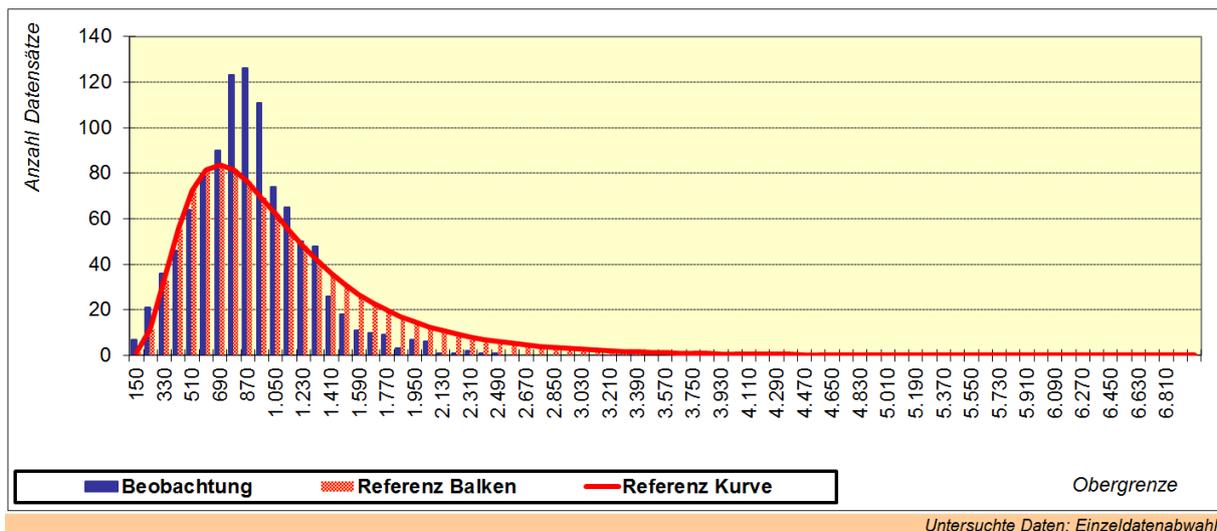
Folgende Hinweise sind für die Anwendung der Funktion „Referenzsumme“ in der SRP-Vorlage Stochastik wichtig:

- ➔ Alle auf dem Tabellenblatt „Daten“ vorgenommenen **generellen Dateneinschränkungen** (s. den Abschnitt „zweistufige Eingrenzungsmöglichkeit“ unter [Pkt. 4.3.2.1.](#)) wirken auch auf die Berechnung des erwarteten Gesamtabsatzes. In der beispielhaften Abb. weisen die **rot gefärbte Überschrift und der Hinweis** darauf hin, dass ein Datensatz zuvor ausgewählt wurde, weil er sich als außergewöhnliche Tageskasseneinnahme einer pauschal abgerechneten Veranstaltung herausgestellt hat.
- ➔ Die Referenzsumme arbeitet grundsätzlich auf Grundlage der angepassten Logarithmischen Normalverteilung (s. [Pkt. 4.3.4.3.](#)). Bei einem vorliegenden Spitzeneingriff ist der Mittelwert der untersuchten Tageskasseneinnahmen ebenfalls betroffen⁴²⁵, so dass eine Hochrechnung ohne Ausrichtung der (Punkt-)Erwartung an denjenigen Daten mit einem geringeren Verkürzungsrisiko keinen Sinn ergeben würde.
- ➔ Weil rechtsschiefe Logarithmische Normalverteilungen weit nach rechts in hohe Wertbereiche auslaufen, besteht die Möglichkeit, einen **Maximalwert** einzugeben. Dadurch werden alle noch nicht „verbrauchten“ Tage mit dem Maximalwert angesetzt. Sollten Tageskasseneinnahmen oberhalb des optionalen Maximums liegen, werden diese nicht nach unten korrigiert.

⁴²⁴ Um die unterschiedlichen asymmetrischen Steigungen in den Abschnitten der Logarithmischen Normalverteilung zu berücksichtigen, wird der mittlere Wertinhalt jeder Schicht mit der gemittelten Häufigkeit aus dem betreffenden und dem vorherigen Intervall multipliziert.

⁴²⁵ Da bei einer Einnahmenverkürzung die Summe der Einnahmen reduziert wird, sich die Anzahl der Tage aber i.d.R. nicht (in gleichem Maß) durch komplettes Streichen verringert, wird der Mittelwert als $\frac{\text{Summe}}{\text{Anzahl}}$ durch die Manipulation ebenfalls verkleinert.

Die Funktion „Referenzsumme“ dient dazu, die großen Auswirkungen von Spitzeneingriffen klarzustellen. Für eine **Schätzung** im Besteuerungsverfahren (s. [Pkt. 4.8.](#)) ist diese Hochrechnung i.d.R. nur bedingt geeignet, weil „[d]ie Schätzungsungewissheit [...] nicht dazu führen [darf], nur den Betrag anzunehmen, der auch im ungünstigsten Fall als sicher vereinbart angesehen werden kann“⁴²⁶. Vielmehr „ist das FA im allgemeinen nicht nur berechtigt, sondern verpflichtet, in dem gegebenen Schätzungsrahmen an die oberste Grenze zu gehen.“⁴²⁷ Dennoch kann die Auswertung in geeigneten Fällen – speziell auch für das Strafverfahren (s. [Pkt. 4.9.](#)) – zur Ermittlung einer sachgerechten Schätzungshöhe beitragen.



Dazu soll das Tabellenblatt „Referenzdiagramm“ der SRP-Vorlage Stochastik eine weitere Absicherung anbieten: Unter linearer Werteklassifizierung auf der X-Achse kann kontrolliert werden, ob die betriebsnah abgeleitete Logarithmische Normalverteilung den wichtigen Bereich mit stark ansteigenden Häufigkeiten der niedrigeren Tageskasseneinnahmen passend wiedergibt. Im Beispieldiagramm werden die Verhältnisse in der Zone „> € 330 bis € 600“⁴²⁸ optimal getroffen. Auch in dieser Auswertung wird der Hinweis auf vorgenommene Einschränkungen der Daten ausgegeben.

4.3.5. Beweiswert von (musterbasierten) Häufigkeitsanalysen

Gesamtbildprüfungen (Verprobungen) sollen die Vergangenheit anhand von Unterlagen, Auskünften und Eindrücken wiederherstellen. Da Informationen zur Vergangenheit niemals vollständig und objektiv sind, ergibt sich grundsätzlich eine **Sachverhaltsbeschreibung mit Unsicherheiten**. Bei Gesamtbildprüfungen ist diese Unsicherheit besonders groß,

⁴²⁶ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 162 AO, Rz. 44.

⁴²⁷ BFH v. 9.3.1967, BStBl. III 1967, 349, Leitsatz.

⁴²⁸ Die X-Achsenbeschriftung benennt nur die Obergrenzen jedes zweiten Intervalls.

weil aufgrund der Komplexität der entscheidungsrelevanten Vorgänge keine einfachen Ursache-Wirkungs-Beziehungen vorliegen können⁴²⁹.

Um dennoch in Gerichtsverfahren und Prüfungen Urteile bzw. ausreichend sichere Beurteilungen zu ermöglichen, werden an den juristischen Beweis geringere Anforderungen gestellt als an den naturwissenschaftlichen. Für den Bereich des Besteuerungsverfahrens sieht das Gesetz auf Seiten der Finanzverwaltung in **§ 158 AO (Verifikationsprinzip)** eine abermals abgesenkte erforderliche Aussagegewissheit vor (s. [Pkt. 4.6.](#)), denn andernfalls ließe sich die (annähernde) Gleichbehandlung aller Stpfl. gem. Art. 3 GG in einem Massenverfahren nicht umsetzen⁴³⁰.

Die vom BFH⁴³¹ zur Schätzung – und der Argumentation folgend ebenso zur Verprobung – bevorzugten Methoden „Vermögenszuwachs-“, „Geldverkehrsrechnung“ und „Kalkulation“ setzen bereits bei ordnungsmäßigen Besteuerungsunterlagen auf unvollständigen Ausgangsinformationen auf, denn bspw. private Geldzu- und -abflüsse bzw. der Warenverderb müssen nicht aufgezeichnet werden. Zudem ist das Risiko ordnungswidriger Besteuerungsunterlagen in den betreffenden Fällen mit überwiegend Bareinnahmen besonders hoch.⁴³² Mindestens unter diesen Umständen kann das Modell „Schwarzeinnahmen zur Bezahlung von Schwarzlöhnen“ mit den Mitteln einer Betriebsprüfung häufig nicht ausgeschlossen werden. I.R. von Kalkulationen liegen oft keine vollständigen und vertrauenswürdigen Informationen zu Verkaufspreisen, Sonderverkaufspreisen, Anwendungsquoten von Sonderverkaufspreisen, Wareneinkäufen etc. vor. In der Konsequenz muss der Sollumsatz auf unsicheren Zahlen rekonstruiert werden, so dass – trotz scheinbar eindeutiger Zielgröße – regelmäßig eine beträchtliche Unsicherheit verbleibt.

⁴²⁹ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15: „Unser Dilemma im Umgang mit der Komplexität unserer Welt lässt sich darauf zurückführen, dass wir wohl darin ausgebildet wurden einfache logische Schlüsse zu ziehen und nahe liegende Ursache-Wirkungs-Beziehungen zu definieren. [...] Simple Ursache-Wirkungs-Beziehungen gibt es nur in der Theorie, nicht in der Wirklichkeit.“

⁴³⁰ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

⁴³¹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 63 u. 64 : „Allein die Ergebnisse eines Zeitreihenvergleichs lassen aufgrund der dieser Verprobungsmethode innewohnenden methodenbedingten Unsicherheiten aber noch keinen sicheren Schluss auf das Vorliegen und den Umfang auch materieller Unrichtigkeiten der Buchführung zu. In diesen Fällen sind andere Schätzungsmethoden, die auf betriebsinternen Daten aufbauen oder in anderer Weise die individuellen Verhältnisse des jeweiligen Steuerpflichtigen berücksichtigen (z.B. Vermögenszuwachs- oder Geldverkehrsrechnung, Aufschlagkalkulation) grundsätzlich vorrangig heranzuziehen.“

⁴³² Teutemacher, „Handbuch zur Kassenführung – Praxishandbuch für die rechtssichere Umsetzung“, NWB Verlag, 2015, 2: „Betriebsprüfungen bei Betrieben, die überwiegend Bargeschäfte tätigen, führen in 95 % der Fälle zu dem Ergebnis, dass die Kassenführung verworfen wird, weil die vorgelegten Bücher bzw. Aufzeichnungen sowie Daten nicht den Ordnungsvorschriften entsprechen und damit die sachliche Richtigkeit in Zweifel gezogen wird“ mit Verweis auf Kalischke, „Betriebe bangen vor Steuerprüfern – Beamte werfen immer mehr Kassenbuchführungen“, Westfälische Nachrichten v. 15.3.2013.

Sorgfältige Ermittlungen der Finanzbehörde⁴³³ sollen die Unsicherheiten soweit möglich begrenzen. Nach dem Verifikationsprinzip haben aber auch die Mitwirkungsbereitschaft und -qualität der Stpfl. eine wesentliche Auswirkung auf das erforderliche Beweismaß für das Auslösen der Schätzungspflicht.⁴³⁴ Bei manipulationsorientierten Musteranalysen bedeutet das sorgfältige⁴³⁵ Erarbeiten eines Gesamtbilds, einem ausgezählten Häufigkeitsmuster (i.d.R. aus Tageskasseneinnahmen) eine betriebsspezifische (Punkt-)Erwartung gegenüberzustellen, die Übereinstimmung mit einem geeigneten Wahrscheinlichkeitstest zu prüfen, damit der vorliegende Datenumfang berücksichtigt wird, und bei Auffälligkeiten die nicht-manipulativen Eingriffsmöglichkeiten weitest möglich abzuklären.

Im Hinblick auf die genannten Kriterien eines sorgfältig erarbeiteten – und damit aussagekräftigen – Gesamtbilds gründet sich die **manipulationsorientierte SRP-Musteranalytik** auf **folgende Grundsätze und Annahmen**:

- ☑ I.R. der SRP werden die (Punkt-)Erwartungen „Gleichverteilung“, „Benford-Gesetz“ und „Logarithmische Normalverteilung“ an den Verhältnissen der zu untersuchenden Daten einzelfallbezogen entwickelt (s. [Pkt. 4.3.1.2.](#)). Das Problem abweichender Erwartungen für die Einerstelle gem. Gleichverteilung bzw. Benford-Gesetz in Werten, die kleiner als „€ 100“ sind, löst die SRP-Vorlage Stochastik automatisch (s. [Pkt. 4.3.2.2.1.](#)). Insbesondere werden die Analyseschritte „(Punkt-)Erwartung“ und „Übereinstimmungsprüfung“ nicht vermischt⁴³⁶, sondern für eine nachvollziehbare Ergebnisbeurteilung konsequent getrennt.

⁴³³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Der Wortlaut des § 158 („Anlass zu beanstanden“) mag dafür sprechen, dass bereits bloße Beweisanzeichen, Zweifel oder Bedenken ausreichen, um die Vermutung auszuräumen oder zu entkräften. Es ist jedoch nicht so, dass die FinBeh. nur irgendwelche Beanstandungen auszusprechen hätte und der Stpfl. dann gehalten wäre zu beweisen, dass die formell richtigen Buchungen sachlich nicht falsch sind. Ein solcher Negativbeweis kann unmöglich verlangt werden. Wenn am Anfang einer Untersuchung Bedenken, Zweifel oder Beweisanzeichen gegen die sachliche Richtigkeit der Buchführung vorliegen, legitimiert (§ 158) und verpflichtet (§ 88) das die Behörde dazu, weitere Nachforschungen zur Sachverhaltsaufklärung anzustellen.“

⁴³⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wie weit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

⁴³⁵ Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz. 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

⁴³⁶ Vgl. z.B.: *Bisler*, „Chi-Quadrat-Test“ und Zeitreihenvergleich: Keine Schätzung bei ordnungsgemäßer Buchführung“, PStR 1/2012, 15: „Der ‚Chi-Quadrat-Test‘ analysiert die Verteilung bestimmter Ziffern und basiert auf der Erkenntnis, dass jeder Mensch Lieblingsziffern hat und bei frei erfundenen Zahlen (z.B. nachgetragenen Tageseinnahmen oder -ausgaben) häufiger verwendet als andere.“ Tatsächlich ist der χ^2 -Test (Chi-Quadrat-Test) ein allgemeiner Wahrscheinlichkeitstest, der mehrklassige Verteilungen unter Berücksichtigung der Er-

- ☑ Die Übereinstimmungsprüfung der beobachteten Häufigkeitsmuster mit den (Punkt-)Erwartungen erfolgt bei der indiziellen Suche nach Manipulationsspuren in der SRP immer durch einen geeigneten Wahrscheinlichkeitstest: für mehrklassige⁴³⁷ Zifferanalysen und die Verteilungsanalyse mit dem χ^2 -Test (Chi-Quadrat-Test) (s. *das Stichwort „Chi-Quadrat-Test“ jeweils unter [Pkt. 4.3.2.1.](#) u. [Pkt. 4.3.4.2.](#)*) und für Zifferanalysen aus der Perspektive einer Einzelziffer mit der Binomialverteilung (s. *das Stichwort „Binomialverteilung“ jeweils unter [Pkt. 4.3.2.1.](#) u. [Pkt. 4.3.2.2.3.](#)*). Dadurch wird die Anzahl untersuchter Datensätze im Übereinstimmungsergebnis berücksichtigt, wobei die Empfehlung zur Mindestdatenmenge für den χ^2 -Test (Chi-Quadrat-Test) von der SRP-Vorlage Stochastik automatisch beachtet wird (s. [Pkt. 4.3.1.3.](#)).
- ☑ Das mögliche Vorliegen nicht-manipulativer Ursachen für auffällige Häufigkeitsmuster (s. *den Abschnitt „musterbasierte Prüfungsansätze zur Aufdeckung von Manipulationen“ unter [Pkt. 4.3.](#)*) wird sowohl bei den SRP-Zifferanalysen als auch bei der SRP-Verteilungsanalyse selbständig untersucht. Dazu reichen bei Einerstellenauffälligkeiten einfache mathematische Überlegungen (s. [Pkt. 4.3.2.2.3.](#)). Bei Zweitstellenauffälligkeiten stellt die SRP-Vorlage Stochastik zielführende Funktionen zur Verfügung (s. [Pkt. 4.3.2.3.3.](#)). Vorbereitende Maßnahmen (s. [Pkt. 4.3.3.3.](#)) und eine gezielte Nachbetrachtung sichern ein mit der Verteilungsanalyse beobachtetes typisches Manipulationsmuster ab. Sollten darüber hinaus konkrete Einlassungen der Stpfl.⁴³⁸ vorliegen, können diese interaktiv in der SRP-Vorlage Stochastik berücksichtigt werden wie z.B. beim Ausschluss eines nicht vergleichbaren Sondersachverhalts (s. *die Abb. zum Aufzählungspkt. „Hilfsminimum und Hilfsmaximum“ unter [Pkt. 4.3.3.1.](#)*).

Die hohe Aussagegewissheit einer Plausibilitätsprüfung i.S. des § 158 AO mit SRP ergibt sich nicht nur aus den einzelnen aufgelisteten Qualitätsmerkmalen, sondern aus der SRP-Musteranalytik, die alle Vorteile über ein systematisches Prüfungsschema (s. *den Abschnitt „Prüfungsschema zur Vermeidung von Fehldeutungen“ unter [Pkt. 4.3.](#)*) verbindet.

kenntnisse der Wahrscheinlichkeitslehre auf Übereinstimmung prüft. Eine (zwingende) Verbindung zur Gleichverteilung von endständigen Ziffern in Wirtschaftsdaten besteht nicht. Vielmehr vermischt die Aussage die zwei Schritte „Erwartung“ und „Übereinstimmungsprüfung“ i.R. der musterbasierten Analyse und sorgt somit für Verständnisprobleme und Fehleinschätzungen.

⁴³⁷ Üblicherweise stellen Zifferanalysen zunächst die Beobachtung und (Punkt-)Erwartung für alle möglichen Ziffern gegenüber, was z.B. für die Einerstelle i.d.R. alle zehn Ziffern bedeutet. Somit werden die Datensätze nach zehn Ziffern, also in zehn Klassen ausgezählt. Soll die Untersuchung im Hinblick auf eine einzelne zu häufig oder selten vorkommende Ziffer vertieft werden, resultieren daraus die möglichen Ergebnisse „Ziffer x“ und „alle anderen Ziffern“, also nur noch zwei Klassen.

⁴³⁸ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit [...] Dabei ist allerdings die sphärenorientierte Mitwirkungspflicht des Stpfl. (s. § 88 Rz. 3 f.; § 90 Rz. 20) zu beachten. Mathematisch eindeutige Unschlüssigkeiten der Buchführungsergebnisse muss er durch betriebspezifische Besonderheiten schlüssig erklären“.

In der bisherigen **Rechtsprechung** der Finanzgerichte werden manipulationsorientierte Musteranalysen nur als zusätzliches Indiz anerkannt.⁴³⁹ Derweil diese Methoden auf wissenschaftlichen Grundlagen basieren⁴⁴⁰ und z.B. zur Aufdeckung von Wahlfälschungen oder Waffenhandel eingesetzt werden⁴⁴¹, ist diese Zurückhaltung nur damit zu erklären, dass eine unsachgemäße Anwendung befürchtet wird⁴⁴². Das SRP-Prüfungsschema zur Vermeidung von Fehldeutungen gewährleistet einen hohen Anwendungsstandard und nachvollziehbare Ergebnisse. Sollten trotz selbständiger Bemühungen und interaktiver Optionen zur Berücksichtigung konkreter Erklärungsversuche⁴⁴³ mathematisch eindeutige (signifikante) Muster auffälligkeiten bestehen bleiben, erreichen diese (annähernd) die indizielle Aussageklarheit, die der Wahrscheinlichkeitstest als Nichtübereinstimmung ausgibt. Damit liegt ein solches Beweisanzeichen gegen die sachliche Richtigkeit weit oberhalb der gesetzlichen Schwelle des § 158 AO von ca. 95 % (s. [Pkt. 4.6.](#)). Deshalb wertet das FG Hamburg jüngst eine

⁴³⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit (s. FG Münster v. 14.8.2003 – 8 V 2651/03 E, U, EFG 2004, 9 m. Anm. *Trossen*; FG Münster v. 10.11.2003 – 6 V 4562/03 E, U, EFG 2004, 236 [237]; FG Münster v. 7.12.2005 – 1 K 6384/03 E, EFG 2006, 652 [653] m. Anm. *Zimmermann* [Fahrtbuch]; FG Rh.-Pf. v. 24.8.2011 – 2 K 1277/10, EFG 2012, 10 [11] m. Anm. *Kühnen*; *Sosna*, St-Bp. 2000, 41 [Teil I], 68 [Teil II]; *Meyer*, DStR 2005, 2114; *Watrin/Stuffert*, DB 2006, 1748; *Huber*, StBp. 2009, 121 [124 ff.]“.

⁴⁴⁰ *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25: „Erwähnenswert ist in diesem Zusammenhang eine Untersuchung der Universität Münster (Institut für Wirtschaftsinformatik) über die Aussagekraft des Chi-Quadrat-Tests. [...] Diese Untersuchung dokumentiert deutlich, dass diese statistische-mathematische Prüfungsmethode auf wissenschaftlich gesicherten Grundlagen beruht.“

⁴⁴¹ Vgl. z.B.: Handelsblatt online v. 8.7.2012: „Volkswirte auf Verbrecherjagd: Ob illegaler Waffenhandel oder Steuerhinterziehung – Ökonomen stöbern immer häufiger Missetäter auf. Die forensische Ökonomie ist ein boomender Forschungszweig. Ermittlungsbehörden nutzen die Erkenntnisse der Forscher“, nachzulesen unter: <https://www.handelsblatt.com/politik/konjunktur/oekonomie/nachrichten/forensische-oekonomie-volkswirte-auf-verbrecherjagd/6838146.html> oder Handelsblatt online v. 30.11.2009: „Auf der Jagd nach Zahlen-Fälschern: Mit einem lange vergessenen mathematischen Gesetz wollen Ökonomen Kollegen überführen, die in wissenschaftlichen Studien Daten manipuliert haben. Denn längst nicht alle Wirtschaftswissenschaftler sind ehrlich – und die Anreize für Betrüger sind groß“, nachzulesen unter: <https://www.handelsblatt.com/politik/konjunktur/oekonomie/nachrichten/oekonomen-auf-der-jagd-nach-zahlen-faelschern/3315384.html>.

⁴⁴² Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz. 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

⁴⁴³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 21a: „Eine Ziffernanalyse durch einen sog. Chi²-Test besitzt aber allenfalls indiziellen Charakter für eine Unregelmäßigkeit [...] Dabei ist allerdings die sphärenorientierte Mitwirkungspflicht des Stpfl. (s. § 88 Rz. 3 f.; § 90 Rz. 20) zu beachten. Mathematisch eindeutige Unschlüssigkeiten der Buchführungsergebnisse muss er durch betriebspezifische Besonderheiten schlüssig erklären“.

stochastisch eindeutige Musterauffälligkeit zu Recht als eigenständiges Indiz – wenn auch unter formellen Mängeln.⁴⁴⁴

4.4. Gesamtergebnis einer plausibilitätsorientierten SRP

Nach einer (plausibilitätsorientierten) SRP werden Schlussfolgerungen gem. § 158 AO ggf. i.V. mit § 162 Abs. 2 S. 2 AO auf Basis des **Gesamtbilds** aus

- der Ordnungsmäßigkeitsprüfung
- und den systematisch-mehrperspektivischen Analysen
- sowie – im Fall beobachteter Auffälligkeiten – den selbständigen Aufklärungsmaßnahmen bzw. den eingepflegten Erklärungen der Stpfl. abgeleitet.

Die systematisch-mehrperspektivische Kombination *geeigneter* regel- und musterbasierter Datenverbildlichungen stellt – wie bei der dreidimensionalen Bildgebung der medizinischen Diagnostik – sicher, dass mögliche gesamtbildrelevante Sondersachverhalte nicht übersehen werden (s. *das Beispiel zu einer Zeitreihenauffälligkeit im Abschnitt „Prüfungsnetz“ unter Pkt. 2.3.*). Dadurch können sie zur sachgerechten Beurteilung interaktiv (s. *Pkt. 4.2.9.*) auf Normalverhältnisse korrigiert werden – bspw. bei wichtigen einseitigen Preiseffekten mittels Preisindizes (s. *das entsprechende Beispiel unter Pkt. 4.2.5.1.*). Auch Auffälligkeiten aus manipulationsorientierten Musteranalysen werden gezielt auf nicht-manipulative Ursachen hinterfragt (s. *Pkt. 4.3.2.2.3., Pkt. 4.3.2.3.3. und Pkt. 4.3.4.5.*).

Im Ergebnis ist bereits jedes **einzelne verbleibende Beweisanzeichen** gegen die sachliche Richtigkeit der untersuchten Besteuerungsgrundlagen soweit abgesichert worden, wie es die Datenqualität und Aufklärungsmitwirkung⁴⁴⁵ im Einzelfall zulässt.

Nach ständigen Prüfungserfahrungen besteht ein gem. § 158 AO zu beurteilendes SRP-Gesamtbild – insbesondere bei Fällen mit überwiegend Bargeschäften – jedoch aus **mehre-
ren Indizien gegen die sachliche Richtigkeit.**

Die folgende Abb.⁴⁴⁶ führt die (wichtigsten) Ergebnisse einer typischen Beanstandung der Einnahmenvollständigkeit infolge einer SRP zusammen. Anhand der knappen⁴⁴⁷ Aufzäh-

⁴⁴⁴ Ur. v. 5.3.2018 – 3 K 205/15, Rz. 66: „Schließlich hat die stochastische Analyse der erklärten Erlöse ergeben, dass die erklärten Zahlen der Samstage mit an Sicherheit grenzender Wahrscheinlichkeit nicht zufällig entstanden, sondern manipuliert worden sind.“

⁴⁴⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „§ 162 II i.V. mit § 158 orientieren das Beweismaß an der Sphärenverantwortlichkeit [...] Das Beweismaß hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

lungspunkte zu festgestellten Ordnungswidrigkeiten in den drei Aufzeichnungsebenen „Einzelerfassung mit elektronischem Kassensystem“, „tägliche Zusammenfassung im Kassenbuch“ und „Buchführung“ lässt sich eindeutig schlussfolgern, dass keine ausreichenden Maßnahmen zur Absicherung der manipulationsanfälligen Kassendaten gegen Eingriffe erfolgt sind (s. den Abschnitt „Gewissheit der Originalität“ unter [Pkt. 4.1.](#)). Wegen der Schwere⁴⁴⁸ dieser Ordnungsmängel ist mangels inhaltlicher Verifizierbarkeit der sachlichen Richtigkeit „systembedingt keine Gewähr mehr für die Vollständigkeit der Erfassung der Bareinnahmen“⁴⁴⁹ herstellbar. Die „Beweiskraft der Buchführung (und Aufzeichnungen)“ ist damit bereits nach § 158 AO erschüttert.

<u>Ordnungsmäßigkeit</u>	<u>SRP-Ergebnisse</u>
<p>§ Erfassung der Bareinnahmen mit einem (speicherfähigen) elektronischen Kassensystem aus Terminal und Ordermen</p> <p>➔ keine Einstellungsprotokolle (Zusammenführung aller Ordermen?)</p>	<p>⚡ Unschlüssige Entwicklung des Warenflusses (z. T. fehlende Einkaufsreaktionen auf den Absatz)</p> <p>➔ Alltagseffekte beim robusten SRP-Modell als Ursache äußerst unwahrscheinlich</p> <p>➔ keine Ausnahmevorgänge auffindbar bzw. nachgewiesen</p>
<p>§ Betriebsinternes Excel-Kassenbuch mit täglicher Eintragung der Bareinnahmen, Ausgaben, Entnahmen und Einlagen sowie fortlaufender Bestandsberechnung</p> <p>➔ keine Sicherheitslösung (technisch/organisatorisch) zur Unterbindung von nachträglichen spurlosen Änderungen</p>	<p>⚡ Durchschnittliche Wochentagserlöse mit Anstieg um 20 % <i>nur</i> an den Freitagen</p> <p>➔ Zufallsentwicklung äußerst unwahrscheinlich</p> <p>➔ kein Nachweis einer regelmäßigen absatzerhöhenden Freitagsaktion</p>
<p>§ Weitergabe der Kassenberichte und Belege i.d.R. einmal monatlich zur Buchführung an das Steuerbüro</p> <p>➔ keine sofortige Festschreibung der Kassenaufzeichnungen (<i>frühestens mit Buchung im Steuerbüro</i>)</p>	<p>⚡ Verteilungsanalyse der Freitags-Bareinnahmen mit typischem Bild des Spitzeneingriffs</p> <p>➔ mathematisch eindeutige (signifikante) Abweichung von der (betriebsspezifischen) Logarithmischen Normalverteilung trotz geringer Datenanzahl</p> <p>➔ keine Doppel- bzw. Mehrfachverteilung und keine Kapazitätsgrenze erkennbar</p>

Aus den auf der rechten Seite aufgelisteten auffälligen Beobachtungen zu den betreffenden Besteuerungsdaten „Einnahmen“ und „Warenfluss“ lässt sich außerdem entnehmen, dass deren mangelhafte Nachprüfbarkeit offenbar den üblichen⁴⁵⁰ Zusammenhang zu wesentli-

⁴⁴⁶ Quelle (überarbeitet): *Webel/Wähnert*, „Der Verdacht einer Steuerstraftat während der Außenprüfung – § 10 BpO im Kontext der Digitalisierung“, NWB 44/2014, 3324, 3332.

⁴⁴⁷ In der Prüfungspraxis sollte die Darlegung von Ordnungswidrigkeiten ausführlicher erfolgen (s. *das Beispiel im Abschnitt „Ordnungsmängel ausführlich dokumentieren“ unter Pkt. 4.1.*).

⁴⁴⁸ Vgl. *Nöcker (RiBFH)*, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3550: „Liegen keine materiellen Buchführungsmängel vor, entscheidet die Gewichtigkeit der formellen Buchführungsmängel darüber, ob eine Schätzungsbefugnis des Finanzamts vorliegt.“

⁴⁴⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

⁴⁵⁰ Vgl. z.B.: *Tipke/Kruse*, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 24: „Allerdings entspricht es der Erfahrung, dass eine formell ordnungswidrige Buchführung i.d.R. auch sachliche Mängel aufweist.“

chen inhaltlichen Unrichtigkeiten bedeutet. Bereits für jedes einzelne der aufgeführten Indizien gegen die sachliche Richtigkeit erreicht die SRP mit ihren speziellen Merkmalen eine hohe Aussagegewissheit:

1. das robuste Zeitreihenmodell (s. die Gegenüberstellung gängiger Zeitreihenmodelle im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) i.V. mit der SRP-Zeitreihensystematik zum Auffinden und Neutralisieren von möglichen Störeinflüssen (s. [Pkt. 4.2.2.](#)),
2. die vielfältigen und variablen internen Vergleichsmöglichkeiten der untersuchten Datenreihen (s. [Pkt. 4.2.7.](#)) mit interaktiven Korrekturoptionen (s. [Pkt. 4.2.9.](#)) sowie
3. die qualifizierte Beurteilung von manipulationsorientierten Musteranalysen mit Hilfe von Wahrscheinlichkeitstests (s. [Pkt. 4.3.1.3.](#)) zuzüglich der gezielten Abklärung möglicher nicht-manipulativer Ursachen (s. [Pkt. 4.3.4.5.](#)).

Mindestens genauso entscheidend für die bessere Aussageklarheit der systematisch-mehrperspektivischen SRP ist zudem die gegenseitige Verstärkung der einzelnen Indizien innerhalb ihrer **Indiziengesamtheit**, wozu der BGH grundsätzlich feststellt: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“⁴⁵¹ Die beweisverstärkende Wirkung mehrerer Indizien ist mathematisch logisch: Ein Indiz ist eine Feststellung zur möglichen Täterschaft mit einer i.d.R. nicht genau bekannten aber höheren Wahrscheinlichkeit. Liegen zwei (unabhängige⁴⁵²) Indizien vor, welche dieselbe Täterschaft vermuten lassen, ermittelt sich die Gesamtrestunsicherheit mit: Restunsicherheit 1⁴⁵³ x Restunsicherheit 2. Im Fall zweier unabhängiger Indizien von jeweils ca. 90 % Gewissheit reduziert sich damit die Restunsicherheit auf 10 % x 10 % = 1 %, und somit steigert sich die Gesamtklarheit (exponentiell) auf 99 %. Bei (möglicherweise) abhängigen Beweisanzeichen ist der verstärkende Effekt weniger extrem aber immer noch entscheidungserheblich.⁴⁵⁴

⁴⁵¹ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20.

⁴⁵² Die einfache Berechnung von Gesamtwahrscheinlichkeiten unabhängiger Ereignisse ist aus der Spieltheorie bekannt. Soll z.B. beim Werfen eines Würfels die Möglichkeit, mehrere Male in Folge eine „6“ zu werfen, ermittelt werden, sind die Einzelwahrscheinlichkeiten zu multiplizieren. Für bspw. dreimal „6“ nacheinander ergibt sich somit: $\frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216}$, also ein ziemlich unwahrscheinliches Ergebnis (ca. 0,5 %).

⁴⁵³ Restunsicherheit Indiz 1 = 100 % – Indizien-gewissheit 1.

⁴⁵⁴ Die Ausführungen der Gutachter Köller/Nissen/Rieß/Sadorf („Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigen-gutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004, 22) beziehen sich auf die Bayes-Logik zur Verknüpfung abhängiger Wahrscheinlichkeiten (Beweisanzeichen) in Gerichtsverfahren.

Mit der gemeinsamen Bewertung von formellen und materiellen Mängeln in Besteuerungsaufzeichnungen wird diese weitreichende Folge von Indiziengesamtheiten auch in der Drei-Stufen-Theorie des BFH⁴⁵⁵ verarbeitet. Indes sind die Auswirkungen beim Zusammenfallen von zwei oder mehreren beobachteten Unschlüssigkeiten auf ein Gesamtbild i.S. des § 158 AO genauso gravierend. In dem oberhalb abgebildeten Beispielsfall zu einem SRP-Gesamtbild kann eine Abhängigkeit der unschlüssigen Beobachtungen zu den Freitagserlösen bzw. -bareinnahmen nicht ausgeschlossen werden. Demnach würde keine vollständige multiplikative Beweisverstärkung resultieren. Unzweifelhaft ist jedoch, dass die implausiblen Warenflussverhältnisse unabhängig von den beiden anderen Indizien sind. Im Ergebnis reduzieren sich die Restunsicherheiten der Plausibilitätsbeurteilung einmal (exponentiell) stark und einmal etwas weniger deutlich, also insgesamt wesentlich. Auch in der Wirtschaftsprüfung werden diese Vorteile der systematisch-mehrperspektivischen Prüfungsweise der SRP inzwischen erkannt.⁴⁵⁶

4.5. SRP im Verhältnis zu anderen Prüfungsmethoden

Die SRP ist ein komplexes Prüfungsnetz zur schnellen Orientierung in größeren Datenmengen. Sie verbindet dazu zwei Wege der Datenvisualisierung mit regel- und musterbasierten Ansätzen inkl. zielführender Anpassungsfunktionen, weil diese die Stärken der menschlichen Wahrnehmung nutzen (s. mit einem Beispiel dazu den Abschnitt "Visualisierung von Steuerungsdaten" unter [Pkt. 2.](#)). Trotzdem werden dadurch die vorgeschaltete Formalprüfung oder **andere Methoden** wie z.B. Vermögenszuwachs- bzw. Geldverkehrsrechnungen oder Kalkulationen **nicht überflüssig**.

Am besten können die Vorteile der SRP ausgenutzt werden, indem **alle zur Verfügung stehenden Prüfungsmittel** funktionsgerecht sowie nach Effektivitäts- und Effizienzerwägungen⁴⁵⁷ in die jeweilige Prüfung **eingebunden** werden:

- ☞ Die **Formalprüfung** von Besteuerungsunterlagen auf Ordnungsmäßigkeit ist zu Beginn einer Außenprüfung grundsätzlich erforderlich (s. [Pkt. 4.1.](#)). Elektronische Daten

⁴⁵⁵ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 62 ff.

⁴⁵⁶ Internetseite der Wirtschaftsprüfungs-/Steuerberatungsgesellschaft Rödl & Partner GmbH, Veröffentlichung v. 2.5.2018: „Vorteile der Summarischen Risikoprüfung als digitale Prüfungsmethode – Schnell gelesen: 1.) Digitale Prüfungsmethoden bringen Vorteile mit sich und ermöglichen ein mehrperspektivisch erarbeitetes Gesamtbild. 2.) Durch eine sinnvolle Verknüpfung vieler Datenverbildlichungen lässt sich eine Schlussfolgerung im Sinne des § 158 AO ableiten.“

⁴⁵⁷ § 7 BpO: „Prüfungsgrundsätze: Die Außenprüfung ist auf das Wesentliche abzustellen. Ihre Dauer ist auf das notwendige Maß zu beschränken. Sie hat sich in erster Linie auf solche Sachverhalte zu erstrecken, die zu endgültigen Steuerausfällen oder Steuererstattungen oder -vergütungen oder zu nicht unbedeutenden Gewinnverlagerungen führen können.“

lassen ohne spezielle Merkmale bzw. Vorkehrungen kaum seriöse Aussagen über ihre Vollständigkeit und Unverändertheit zu. Auch Gesamtbildprüfungen sind unter solchen Umständen nur begrenzt aussagekräftig. Deshalb können gravierende Ordnungsmängel direkt die Amtspflicht der Schätzung auslösen.⁴⁵⁸

 Mit dem Ziel eines sinnvollen Ressourceneinsatzes sollte bei jeder Außenprüfung anschließend die **Erarbeitung eines Überblicks** vorgenommen werden, um gezielt die einzelfallbezogenen Prüffelder festzustellen und ggf. das Gesamtbild nach § 158 AO zu beurteilen. Besonders infolge der zunehmenden Verbreitung digitaler Betriebsverwaltungssysteme mit immer mehr aufgezeichneten Informationen⁴⁵⁹ (s. [Pkt. 2.1.](#)) ist selbst bei kleinen Unternehmen eine vollständige Überprüfung aller steuerrelevanten Unterlagen nur mit einem absolut unverhältnismäßigen Zeit- und Arbeitsaufwand möglich. Deshalb sind Übersichtsmethoden, welche betriebsbezogen risikobehaftete Sachverhalte entdecken können, von zentraler Bedeutung für ein funktionierendes Risikomanagement. Dieser zielführende Überblick ist die Hauptfunktion der **SRP** (s. [Pkt. 2.](#)).

 **Vermögenszuwachs- bzw. Geldverkehrsrechnungen und Kalkulationen** sind arbeitsaufwendige Prüfungsmaßnahmen, worauf nur Stpfl. Anspruch haben, die ihren Mitwirkungspflichten nachkommen.⁴⁶⁰ Falls effizientere Übersichtsmittel wie die SRP trotz Mitwirkungsbereitschaft der Stpfl. nicht zu einer ausreichend klaren Beurteilungslage für § 158 AO führen, können sie partiell – und damit unter angemessenem Aufwand – zur Absicherung ergänzt werden.

 Auch **mathematisch-statistische Stichprobenverfahren**, wie sie in der Wirtschaftsprüfung z.B. in Form des "Monetary Unit Sampling (**MUS**)"⁴⁶¹ üblich sind, verfolgen den Gedanken eines risikogerechten Prüfungsumfangs. Die Anwendungsbereiche und Konsequenzen unterscheiden sich jedoch von systematisch-mehrperspektivischen

⁴⁵⁸ Vgl. Nöcker (RiBFH), „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548, 3550: „Liegen keine materiellen Buchführungsmängel vor, entscheidet die Wichtigkeit der formellen Buchführungsmängel darüber, ob eine Schätzungsbefugnis des Finanzamts vorliegt.“

⁴⁵⁹ Vgl. z.B.: Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876, 877: „Die Buchführung hat sich von einer Durchschreibebuchführung hin zu einer umfassenden Datenverarbeitung weiterentwickelt, die elektronische Form ist nunmehr die übliche. [...] In den Unternehmen [...] werden unterschiedlichste betriebliche Aufgaben und Anforderungen in Datenbanken protokolliert oder über Datensysteme gesteuert. Dabei steigt die Anzahl elektronisch überwachter Prozesse stetig.“

⁴⁶⁰ Vgl. z.B.: BFH v. 21.2.1990, BFH/NV 1990, 683, Rz. 31: „Der Anregung des Klägers, eine Geldverkehrsrechnung zu erstellen, braucht das FG nicht nachzukommen. Die Geldverkehrsrechnung ist eine arbeitsaufwendige Schätzungsmethode. Die vom Kläger erklärten Buchführungsergebnisse beruhen auf einer formell ordnungswidrigen Buchführung. Daher ist die Anwendung weniger arbeitsaufwendiger Schätzungsmethoden statthaft.“

⁴⁶¹ Vgl. z.B.: Giezek, „Monetary Unit Sampling – Der Einsatz statistischer Verfahren im Rahmen der Jahresabschlussprüfung“, Gabler Verlag, 2011.

Datenvisualisierungen wie der SRP: MUS eignet sich für Niedrigrisikobereiche mit der Gefahr von (punktuellen) Überbewertungen⁴⁶². Aufgefundene Fehler werden i.d.R. nicht nachverfolgt und weiter konkretisiert sondern zum wahrscheinlichsten Fehler der Grundgesamtheit⁴⁶³ hochgerechnet. Eine zielführende Kombination aus MUS und SRP kann durch erhebliche Effizienzgewinne wesentlich zur Gleichmäßigkeit der Besteuerungskontrolle gem. § 85 AO beitragen. Demzufolge ist das Pilotprojekt "Monetary Unit Sampling (MUS)" in Rheinland-Pfalz durch eine Kooperation mit der SRP begleitet worden.⁴⁶⁴

 Demgegenüber ist der medial präsente⁴⁶⁵ Trend einer weitgehenden Automatisierung der Prüfungsprozesse durch „**künstliche Intelligenz (KI)**“, wie er auch in Form multifunktionaler Werkzeuge von den Softwareherstellern beworben wird⁴⁶⁶, kritischer zu sehen. Die selbstlernende Anpassung von Analysealgorithmen ist kaum nachprüfbar und grenzt den Menschen aus. Da jedoch die Problemlösungen des menschlichen Gehirns chaotisch assoziativ und somit anders erfolgen als die geordneter Datenbanken, bleibt auf diesem Weg viel Prüfungskompetenz ungenutzt. Dieser Mangel wird in den Führungsebenen der Wirtschaft zumindest „gefühl“ wahrgenommen.⁴⁶⁷

4.6. Welches Beweismaß fordert § 158 AO?

Das **Verifikationsprinzip** soll die Verfassungsmäßigkeit der Steuerbelastung dadurch sicherstellen, dass eine gleich- und gesetzmäßige Besteuerung tatsächlich umgesetzt werden kann.⁴⁶⁸ Dafür stellt es mit § 158 AO zwei Voraussetzungen an die „Beweiskraft der Buchführung (und Aufzeichnungen)“:

1. deren Ordnungsmäßigkeit gem. §§ 140 bis 148 AO und

⁴⁶² Für eine mathematisch-statistische Stichprobenprüfung im Hinblick auf mögliche Unterbewertungen muss das Auswahlverfahren des Monetary Unit Sampling (MUS) abgeändert werden.

⁴⁶³ Most Likely Error (MLE).

⁴⁶⁴ Bundesfinanzministerium, Bundesfinanzakademie, Quartals-Info I/2018, „Podiumsdiskussion: Monetary Unit Sampling (MUS) – mathematisch-statistische Stichproben in der Betriebsprüfung“, nachzulesen unter https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Ministerium/Geschaeftsbereich/Bundesfinanzakademie/Aktuelles/BFA_Quartalsinfo_2018_I.html.

⁴⁶⁵ *Kompenhans/Wermelt*, „Big Data – Potenziale für innovative Abschlussprüfungen“, BB 2018, 299: „Die an Systemschnittstellen erforderlichen manuellen Tätigkeiten werden immer mehr durch den Einsatz von Robotic Process Automation ersetzt, Entscheidungsprozesse werden dabei durch den Einsatz programmierter, selbstlernender Entscheidungsheuristiken, die auf Basis künstlicher Intelligenz arbeiten, unterstützt.“

⁴⁶⁶ Nachzulesen unter: <https://blog.audicon.net/audit-analytics/ideav10-visualisierung-und-ermitteln.html>.

⁴⁶⁷ CRN online v. 24.2.2010: „Management-Studie: Manager vertrauen mehr auf ihr Bauchgefühl als auf B[usiness] I[n]telligence“, nachzulesen unter: <https://www.crn.de/software-services/artikel-80249.html>.

⁴⁶⁸ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

2. das Nichtvorhandensein eines „Anlass[es ...], ihre sachliche Richtigkeit zu beanstanden“.

Im Grundsatz löst eine unzureichende Mitwirkung der Stpfl. im Besteuerungsverfahren danach die Amtspflicht der Schätzung aus⁴⁶⁹, weil „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO).“⁴⁷⁰

Mit dem „**Beanstandungsanlass**“ hat der Gesetzgeber für das Besteuerungsverfahren ein besonderes Beweismaß bestimmt. Nach der allgemeinen Gesetzesformulierung und weil die „gesetzliche Vermutung [der sachlichen Richtigkeit von Besteuerungsgrundlagen ...] ihre Wirksamkeit mit der Folge der Schätzungsnotwendigkeit nach § 162 AO [verliert], wenn es nach Verprobung usw. unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt“⁴⁷¹, gilt es für alle Sachverhalte, für welche die Finanzverwaltung die Beweislast trifft. Auf diese Weise wird der Gesetzesauftrag des Verifikationsprinzips umgesetzt. Neben der Verifizierbarkeit im Einzelfall hat die Prüfungsbreite eine fundamentale Bedeutung für die Gleichmäßigkeit der Besteuerung gem. § 85 AO. Ansonsten könnte der intensive Eingriff einer Außenprüfung im Einzelfall an sich bereits als staatliche Willkür empfunden werden.⁴⁷² „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild der Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“⁴⁷³

Das Beweismaß „Beanstandungsanlass“ ist **unabhängig von der Ordnungsmäßigkeit** der Buchführung und Aufzeichnungen, denn es gilt lt. Gesetz ausdrücklich für „[d]ie Buchführung und die Aufzeichnungen des Steuerpflichtigen, die den Vorschriften der §§ 140 bis 148 [AO] entsprechen“. Dennoch senkt ein Zusammentreffen von Ordnungswidrigkeit und materiellen Mängeln das Beweismaß zur sachlichen Verwerfung erklärter Besteuerungsgrundlagen – i.S. von Indiziengesamtheiten⁴⁷⁴ – zusätzlich ab (s. *den Abschnitt „Indiziengesamtheit“ unter [Pkt. 4.4.](#)*).

⁴⁶⁹ § 162 Abs. 2 S. 1 AO: „Zu schätzen ist insbesondere dann, wenn der Steuerpflichtige [...] seine Mitwirkungspflicht nach § 90 Abs. 2 verletzt.“

⁴⁷⁰ BFH v. 15.2.1989, BStBl. II 1989, 462: „[...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

⁴⁷¹ Anwendungserlass zu § 158 AO.

⁴⁷² Tipke, „Die Steuerrechtsordnung“, Verlag Dr. Otto Schmidt, Bd. III, 1407.

⁴⁷³ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5.

⁴⁷⁴ BGH v. 27.4.2010 – 1 StR 454/09, NSTZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die

Eine nähere **Konkretisierung der Aussagegewissheit** „Beanstandungsanlass“ findet sich im Gesetz nicht. Durch die Normierung wird direkt nur klar, dass der Beanstandungsanlass als eigenständiges Beweismaß vom „Beweis“ zu differenzieren ist.⁴⁷⁵ In einer vom Bundeskriminalamt herausgegebenen Arbeit mit dem Untertitel „Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“ plädieren die vier Autoren, die als Gutachter für deutsche Kriminalämter tätig sind/waren, für eine zahlenmäßige Verdeutlichung von Aussagegewissheiten, wie sie in Gerichtsverfahren (und Prüfungen) regelmäßig verwendet werden: „Die numerische Schätzung bewirkt insbesondere eine Vermeidung von Missverständnissen oder möglichen Unterschieden in den scheinbar gleichen Größenordnungen verschiedener Individuen bei identischen verbalen Ausdrücken. [...] Sie erfordert vom Sachverständigen schlicht eine präzisere Einschätzung. Sie erhöht die intersubjektive Beurteilbarkeit dessen, was gemeint ist.“⁴⁷⁶ Ohne eine numerische Konkretisierung ist also nach Gutachtererfahrungen keine Gleichmäßigkeit und Nachvollziehbarkeit von Urteilen gegeben. Auch der Untersuchungsgrundsatz im Besteuerungsverfahren nach § 88 AO (bzw. § 76 FGO) und der Schätzungszusammenhang § 158 i.V. mit § 162 Abs. 2 S. 2 AO können ohne möglichst konkrete Einordnung des gesetzlichen Beweismaßes „Beanstandungsanlass“ nicht einheitlich und nachvollziehbar ausgelegt werden.

Eine erste Orientierung zur Aussagegewissheit „Beanstandungsanlass“ ergibt sich aus der (eindeutigen) **Abgrenzung vom Beweis**. Diese ist lediglich nach unten möglich, weil ein Beweis im besten Fall 100 % sicher ist. Hierbei muss aber zwischen dem naturwissenschaftlichen und dem juristischen Verständnis des Beweises unterschieden werden:

- ! Beweise im naturwissenschaftlichen Sinn verlangen eine strenge Kausalität im Sinn einer zweifelsfreien Ursache-Wirkungs-Beziehung. Durch sorgfältige Beobachtungen der Gegenwart kann diese (nahezu) 100 %-ige Klarheit erreicht werden.
- ! Für einen **Beweis im juristischen Sinn** als volle richterliche Überzeugung muss ein geringerer Maßstab gelten, weil sich die Vergangenheit nicht vollständig und objektiv rekonstruieren lässt. Dokumentationen, Aussagen und beobachtende Rückschlüsse aus der Gegenwart auf die Vergangenheit können immer nur ein Abbild unter Unsicherheit herstellen. Folgerichtig hat der BGH i.R. einer strafrechtlichen Bestimmung des Schuld-

entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

⁴⁷⁵ Hätte der Gesetzgeber mit dem „Beanstandungsanlass“ einen Beweis gemeint, wäre die gesetzliche Formulierung redundant und damit unnötig, denn § 158 AO würde dann ausdrücken, dass bewiesenermaßen falsche Besteuerungsunterlagen keine Beweiskraft hätten.

⁴⁷⁶ Köller/Nissen/Rieß/Sadorf, „Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004, 13.

umfangs „die ermittelte Vertrauensuntergrenze [...] mit einer Wahrscheinlichkeit von 99,5 % [als] zutreffend festgestellt.“⁴⁷⁷ Diese konkretisierende Einordnung des juristischen Beweises steht in Übereinstimmung mit der wissenschaftlichen Bewertung moderner Beweismittel wie DNA- und Fingerabdruckanalysen, welche nach (Meta-)Studien eine geschätzte Restunsicherheit von 0,5-1 % hinterlassen⁴⁷⁸.

Auf Basis der Klarstellung des juristischen Beweises kann die gesetzlich geforderte Gewissheit zur Beanstandung von Besteuerungsgrundlagen durch die Finanzbehörde korrekter und konkreter eingegrenzt werden. Dazu ist die abgebildete tabellarische Übersetzung wichtiger Wahrscheinlichkeitsaussagen aus Gerichtsverfahren in numerische Wahrscheinlichkeiten⁴⁷⁹ eine praktikable und fundierte Hilfe:

Verbale Wahrscheinlichkeiten und deren numerische Zuordnungen

mit indifferenter Wahrscheinlichkeit (non liquet)	1–1/2	50 %
mit leicht überwiegender Wahrscheinlichkeit	1–1/4	75 %
mit überwiegender Wahrscheinlichkeit	1–1/10	90 %
mit hoher Wahrscheinlichkeit	1–1/20	95 %
mit sehr hoher Wahrscheinlichkeit	1–1/100	99 %
mit an Sicherheit grenzender Wahrscheinlichkeit	1–1/10.000	99,99%

Die Autoren begründen die Art der Abstufung mit „der Erkenntnis, dass eine logarithmische Beziehung zwischen Information und Wahrscheinlichkeit besteht“⁴⁸⁰. Aus der höchstrichterlichen Rechtsprechung und gutachterlichen Erfahrungen lässt sich ableiten, dass das gesetzliche Beweismaß „**Beanstandungsanlass**“ mit einer Aussagegewissheit von **ca. 95 %** klar vom Beweis der beiden höchsten Stufen abgegrenzt werden kann.

Die Konkretisierung des Beanstandungsanlasses auf ca. 95 % passt zudem zu **anderen Lebens- und Rechtsbereichen**, in denen Wahrscheinlichkeitsaussagen üblich und notwendig sind: die industrielle Qualitätssicherung⁴⁸¹, die Absicherung von wissenschaftli-

⁴⁷⁷ Urt. v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 26.

⁴⁷⁸ Ausführlich hierzu: *Giezek/Rupprecht/Wähnert*, „Wie sicher muss ein Verprobungsergebnis sein? – Anleihen aus der Kriminalistik zur sinnvollen Einordnung des Beanstandungsanlasses“, BBK 2017, 236.

⁴⁷⁹ *Köller/Nissen/Rieß/Sadorf*, „Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004, 33.

⁴⁸⁰ *Köller/Nissen/Rieß/Sadorf*, „Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004, 32 f. Gemeint ist mit dieser Aussage, dass für eine weitere Steigerung der Klarheit einer beschriebenen Situation zunehmend mehr Informationen notwendig sind.

⁴⁸¹ DIN ISO 2859-1:2014-08: Annahmestichprobenprüfung anhand der Anzahl fehlerhafter Einheiten oder Fehler (Attributprüfung), nachzulesen unter: <https://www.beuth.de/de/norm/din-iso-2859-1/191697831>.

chen Studienergebnissen⁴⁸², die Bewertung großer Datenmengen mit Hilfe von Stichproben⁴⁸³ und die automatische Prozesssteuerung⁴⁸⁴. Aus der mathematischen Sicht des (natürlichen) Streuverhaltens beschreiben die „ca. 95 %“ den Wahrscheinlichkeitsbereich „Mittelwert \pm doppelte Standardabweichung⁴⁸⁵“ für alle Normalverteilungen. Es werden also alle normalen Daten bzw. Situationen zuzüglich der schwachen Ausreißer toleriert und ausschließlich die starken Ausreißer abgelehnt.

Die Auslegung des gesetzlichen Beweismaßes „Beanstandungsanlass“ als Aussagegewissheit „mit an Sicherheit grenzender Wahrscheinlichkeit“ durch den BFH⁴⁸⁶ passt keinesfalls in diesen Kontext. Sowohl im allgemeinen als auch im wissenschaftlichen Sprachgebrauch wird hierunter eine Aussageklarheit minimal unterhalb 100 % verstanden, wie auch die Tabelle „Verbale Wahrscheinlichkeiten und ihre numerischen Zuordnungen“ oberhalb belegt. Nicht einmal juristische Beweise erfüllen diese Anforderung regelmäßig.

Grundsätzlich wäre **Rechtsprechung** am Maßstab „mit an Sicherheit grenzender Wahrscheinlichkeit“ weitgehend unmöglich, weil sich die Vergangenheit in den meisten Fällen nicht mit den begrenzten Informationen bis zu dieser Klarheit wiederherstellen lässt. Absolut unmöglich wäre unter diesem Beweismaß die – nach dem Verifikationsprinzip gewünschte – Beurteilung von Gesamtbildern⁴⁸⁷, weil es zu komplexen Situationen „[s]imple Ursache-Wirkungs-Beziehungen [...] nicht in der Wirklichkeit [gibt]“⁴⁸⁸. Im Ergebnis dürften Verprobungen⁴⁸⁹ in der Prüfung von Besteuerungsgrundlagen keine Bedeutung haben, weil sie das erforderliche Beweismaß generell nicht erreichen können. Tatsächlich finden sich aber zahllose Finanzgerichtsurteile zu Vermögenszuwachs- bzw. Geldverkehrsrechnungen und Kalkulationen, welche aufgrund der unvollständigen und nur bedingt verifizierbaren Ausgangsgrößen *keine* Ergebnisklarheit von 99,99 % erzielen können.

⁴⁸² Vgl. z.B.: Bortz/Schuster, „Statistik für Human- und Sozialwissenschaftler“, Springer-Verlag, 2010, 11.

⁴⁸³ § 241 Abs. 1 HGB: „Bei der Aufstellung des Inventars darf der Bestand der Vermögensgegenstände nach Art, Menge und Wert auch mit Hilfe anerkannter mathematisch-statistischer Methoden auf Grund von Stichproben ermittelt werden“ sowie BGH v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 20 ff.

⁴⁸⁴ Vgl. z.B.: Richter, „Prinzipien der Künstlichen Intelligenz“. Teubner-Verlag, 1992.

⁴⁸⁵ Die Standardabweichung beschreibt die Streuungsintensität einer Datenmenge als vorzeichenneutrale durchschnittliche Abweichung jedes einzelnen Datensatzes zum Mittelwert.

⁴⁸⁶ BFH v. 24.6.1997, BStBl. II 1998, 51: „Nur wenn die Würdigung des Sachverhalts ergibt, daß eine formell ordnungsmäßige Buchführung mit an Sicherheit grenzender Wahrscheinlichkeit sachlich unrichtig ist, kann das Ergebnis der Buchführung ganz oder teilweise verworfen werden.“

⁴⁸⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

⁴⁸⁸ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15.

⁴⁸⁹ Anwendungserlass zu § 158 AO: „Die Vorschrift enthält eine gesetzliche Vermutung. Sie verliert ihre Wirksamkeit mit der Folge der Schätzungsnotwendigkeit nach § 162 AO, wenn es nach Verprobung usw. unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt.“

Trotz dieses unverkennbaren Widerspruchs wird die unzutreffende Auslegung des Beanstandungsanlasses nach § 158 AO als Aussagegewissheit „mit an Sicherheit grenzender Wahrscheinlichkeit“ in der Kommentierung⁴⁹⁰ übernommen. In der Literatur wird hingegen jüngst zur korrekten Einordnung des Beanstandungsanlasses Stellung genommen: „Widerlegen der Richtigkeitsvermutung [...] In der Literatur wird mit guten Argumenten ein Gewissheitsrahmen von 90-95 % für ausreichend gehalten. Das entspricht dem für das finanzgerichtliche Verfahren allgemein geltenden Grundsatz, dass der Richter keine von allen Zweifeln freie Überzeugung anstreben darf, sondern sich in tatsächlich zweifelhaften Fällen vielmehr mit einem für das praktische Leben brauchbaren Grad von Gewissheit begnügen muss.“⁴⁹¹

4.7. Einleitung des Strafverfahrens nach § 10 BpO

Nach § 152 Abs. 2 StPO muss jedes Strafverfolgungsorgan, wozu gem. § 386 Abs. 1 AO auch die Finanzbehörde gehört, beim Vorliegen eines Anfangsverdachts das Strafverfahren einleiten. Das **Legalitätsprinzip** lässt dahingehend keine Abwägung zu, sondern erteilt den Befehl, bei „zureichende[n] tatsächliche[n] Anhaltspunkte[n] für eine Straftat“ einzuschreiten. Die Einleitung eines Steuerstrafverfahrens bewirkt eine Veränderung der rechtlichen Stellung des Steuerpflichtigen als nunmehr Beschuldigter eines Steuerstrafverfahrens.⁴⁹² Einerseits werden dadurch die Ermittlungsmöglichkeiten der Finanzverwaltung um strafprozessuale Maßnahmen erweitert. Andererseits „bestehen [gegenüber dem Besteuerungsverfahren] deutliche Unterschiede im Hinblick auf die Rechte, die dem Beschuldigten zukommen“⁴⁹³. Insbesondere muss er sich nicht selbst belasten, so dass die Einhaltung seiner Mitwirkungspflichten im Besteuerungsverfahren nicht mehr mit Zwangsmitteln durchsetzbar ist. Deshalb führt ein „Verstoß gegen Belehrungspflichten über die Aussagefreiheit des Beschuldigten [...] im Strafprozeß grundsätzlich zu einem Verwertungsverbot der unmittelbar

⁴⁹⁰ Vgl. z.B.: Hübschmann/Hepp/Spitaler, AO/FGO, 242. Lieferung 05.2017, § 158 AO, Rz. 5: „Nach der Rspr. kann eine formell ordnungsmäßige Buchführung ganz oder teilweise verworfen werden, wenn die Würdigung des Sachverhalts mit an Sicherheit grenzender Wahrscheinlichkeit ergibt, dass das Ergebnis der Buchführung sachlich unrichtig ist.“ Gosch, Abgabenordnung/Finanzgerichtsordnung, 140. Lieferung 1.5.2002, § 158 AO, Rz. 11: „Weil eine formell ordnungsmäßige Buchführung die Vermutung der sachlichen Richtigkeit für sich hat, muss die Behörde aber den Nachweis führen, dass das Buchführungsergebnis schlechterdings nicht zutreffen kann, also mit an Sicherheit grenzender Wahrscheinlichkeit materiell unrichtig ist.“ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14: „Die Beanstandungen der sachlichen Richtigkeit durch die FinBeh. müssen mithin so konkret und glaubwürdig (fundiert) sein, dass sie das Ergebnis der Buchführung außerhalb der Wahrscheinlichkeit liegend erscheinen lassen“

⁴⁹¹ Brinkmann, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017, 226 f.

⁴⁹² § 393 Abs. 1 AO: „Die Rechte und Pflichten der Steuerpflichtigen und der Finanzbehörde im Besteuerungsverfahren und im Strafverfahren richten sich nach den für das jeweilige Verfahren geltenden Vorschriften.“

⁴⁹³ Webel, „Steuerfahndung – Steuerstrafverteidigung“, NWB Verlag, 3. Auflage 2016, 462.

hierdurch erlangten Informationen“⁴⁹⁴. Für das Besteuerungsverfahren gilt dies nicht: „Ein Verwertungsverbot im Besteuerungsverfahren bei einer Verletzung des § 393 Abs. 1 Satz 4 AO [...] ergibt sich auch nicht aus allgemeinen verfassungsrechtlichen, insbesondere rechtsstaatlichen Grundsätzen. Vor allem folgt aus der Änderung der Rechtsprechung des BGH, die nunmehr ein strafprozessuales Verwertungsverbot bei unterlassener Belehrung bejaht [...], kein steuerrechtliches Verwertungsverbot.“⁴⁹⁵

Der **Anfangsverdacht** kann sich zum einen auf die Mitwirkungsbereitschaft beziehen, weil „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf“⁴⁹⁶, indem mangels (prüfbarer) Besteuerungsaufzeichnungen auf die Strafverfolgung verzichtet werden müsste. Zum anderen wird im Gesetz eine Gewissheit für die Annahme einer Steuerstraftat festgelegt: „[N]ach kriminalistischer Erfahrung [muss] die Möglichkeit besteh[en], daß eine verfolgbare Straftat vorliegt“⁴⁹⁷. „Trotz der Einbeziehung der ‚kriminalistischen Erfahrung‘ in die Definition des Anfangsverdachts müssen konkrete tatsächliche Erkenntnisse die Basis [...] bilden, die über statistisch gewonnene Erkenntnis[se] bzgl. der Häufigkeit von Straftaten in bestimmten Lebenszusammenhängen und/oder Örtlichkeiten hinausgeh[en].“⁴⁹⁸ In der Konsequenz muss z.B. i.R. einer Betriebsprüfung das Strafverfahren eingeleitet werden, wenn es infolge konkreter Erkenntnisse zum Einzelfall wahrscheinlicher ist, dass eine Straftat vorliegt als dass keine stattgefunden hat.

„Grundsätzlich <i>keine tatsächlichen</i> Anhaltspunkte [...] liegen insbesondere für folgende beispielhaft aufgezählte Fälle vor:	<i>Tatsächliche</i> Anhaltspunkte [...] werden insbesondere in folgenden beispielhaft aufgezählten Fällen gegeben sein:
Bei Aufdecken bloß formeller oder auch kleinerer materieller Buchführungsmängel; [...]	Bei Feststellung schwerwiegender Buchführungsmängel, [...]
Das bloße Durchführen von Kalkulationen und Verprobungen, wie Geldverkehrsrechnungen, Richtsatzverprobungen, Chi ² -Test, Zeit-Reihen-Vergleich usw., [...]	Nach durchgeführter Kalkulation oder Verprobung verbleiben ungeklärte Differenzen von einige[m] Gewicht, [...]

⁴⁹⁴ BGH v. 16.6.2005 – 5 StR 118/05.

⁴⁹⁵ BFH v. 23.1.2002, BStBl. II 2002, 328.

⁴⁹⁶ BFH v. 15.2.1989, BStBl. II 1989, 462: „[...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

⁴⁹⁷ BGH v. 21.4.1988 – III RZ 255/86, NSTZ 1988, 510.

⁴⁹⁸ *Webel*, „Steuerfahndung – Steuerstrafverteidigung“, NWB Verlag, 3. Auflage 2016, 440 mit Verweis auf *Radtke*, in *Radtke/Hohmann*, StPO, § 152 Rn. 19.

Der gleichlautende Erlass der obersten Finanzbehörden der Länder „Verdacht einer Steuerstrafat, Unterrichtung der Bußgeld- und Strafsachenstelle: Die obersten Finanzbehörden der Länder haben zu Anwendungsfragen zu § 10 Abs. 1 BpO Stellung genommen“⁴⁹⁹ konkretisiert den Anfangsverdacht des § 152 Abs. 2 StPO anhand **möglicher Betriebsprüfungserkenntnisse**. Dazu werden die Positionen „keine tatsächlichen Anhaltspunkte“ und „tatsächliche Anhaltspunkte“ für typische Prüfungssituationen hinsichtlich der Ordnungsmäßigkeit und des materiellen Gesamtbilds gegenübergestellt.

In der ergänzenden Klarstellung zu „**schwerwiegende[n] Buchführungsmängel[n]**“ wird auf das „Fehlen von sonst üblichen Belegen“⁵⁰⁰ abgestellt, wodurch die Überprüfung der Vollständigkeit (weitest gehend) verhindert wird. Die Ausführungen *unter Pkt. 4.1.* zur Anwendung der abstrakten Ordnungsvorgaben auf die heutigen, überwiegend digitalen Aufzeichnungs- und Aufbewahrungsweisen gelten gleichermaßen: „Der erkennende Senat vertritt hierzu die Auffassung, dass das Fehlen einer lückenlosen Dokumentation zur Kassenprogrammierung in seinen Auswirkungen auf die Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis dem Fehlen von Tagesendsummenbons bei einer Registrierkasse bzw. dem Fehlen täglicher Protokolle über das Auszählen einer offenen Ladenkasse gleichsteht. In allen drei Fällen lässt der formelle Mangel zwar keinen sicheren Schluss auf die Verkürzung von Einnahmen zu. Gleichwohl gibt es systembedingt keine Gewähr mehr für die Vollständigkeit der Erfassung der Bareinnahmen, ohne dass eine nachträgliche Ergänzung der Dokumentation bzw. eine anderweitige Heilung des Mangels möglich wäre.“⁵⁰¹ Auf die besondere Bedeutung der formellen Ordnungsmäßigkeit von Besteuerungsaufzeichnungen für das Vertrauen in die sachliche Richtigkeit unter den Bedingungen der elektronischen Betriebsverwaltung (s. *Pkt. 2.1.*) weist die AO-Kommentierung ebenfalls hin: „Benutzt der buchführende Stpfl. jedoch eine Manipulationssoftware [...], fehlt es bereits an einer formell ordnungsgemäßen Buchführung, die einen Vertrauensvorschuss verdient [...] verwendet der Stpfl. etwa eine manipulierbare PC-Kasse [...], ist die Beweiskraft einer formell ordnungsgemäßen Buchführung – auch schon vor Inkraft[t]reten des § 146a (ab VZ 2020) – erschüttert“⁵⁰².

Das im Strafrecht geltende Legalitätsprinzip verbietet eine monetäre Bewertung der „**ungeklärten Differenzen von einige[m] Gewicht**“ aus einer Gesamtbildprüfung. Mit den (begrenzten) Mitteln einer Außenprüfung sind zuverlässige Aussagen zum Umfang möglicher Steuerhinterziehungen regelmäßig nicht möglich; hierbei sind die „ausgefeilt ge-

⁴⁹⁹ BMF-Schreiben v. 31.8.2009, BStBl. I 2009, 829.

⁵⁰⁰ BMF-Schreiben v. 31.8.2009, BStBl. I 2009, 829.

⁵⁰¹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 27.

⁵⁰² Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 14 u. 21a.

plante[n] Doppelverkürzungen und/oder [der] Einsatz von Manipulationssoftware, die die Aufdeckungsmöglichkeiten herkömmlicher Prüfungsmethoden minimieren“⁵⁰³ zu bedenken. Deshalb muss der Gradmesser für einen materiellen Anfangsverdacht die Aussagegewissheit sein, dass die vorgelegten Besteuerungsunterlagen inhaltlich nicht richtig sein können. Die Vorteile systematisch-mehrperspektivischer Prüfungsnetze wie der SRP wirken diesbezüglich ebenso:

- ✓ Einzelne Indizien gegen die sachliche Richtigkeit von untersuchten Besteuerungsgrundlagen sind soweit gegen mögliche gesamtbildrelevante Sondersachverhalte bzw. nicht-manipulative Ursachen abgesichert worden, wie es die Datenqualität und Aufklärungsmitwirkung im Einzelfall zulässt (s. den Abschnitt „*einzelne verbleibende Beweisanzzeichen*“ unter [Pkt. 4.4.](#)).
- ✓ I.d.R. besteht das SRP-Gesamtbild aus einer beweisverstärkenden⁵⁰⁴ Gesamtheit von (gravierenden) Ordnungsmängeln und/oder mehreren Indizien (s. den Abschnitt „*mehrerer Indizien gegen die sachliche Richtigkeit*“ unter [Pkt. 4.4.](#)).

In der Literatur⁵⁰⁵ wird die Bedeutung multiindizieller Beanstandungen nach einer SRP für die Einleitung des Strafverfahrens und darüber hinaus festgestellt.

4.8. Schätzung im Besteuerungsverfahren (Quantilsschätzung)

„Mit § 162 AO hat sich der deutsche Gesetzgeber für die Fälle nicht möglicher Ermittlungen [der besteuereungsrelevanten Sachverhalte] für die **Schätzung** entschieden, und das wohl aus gutem Grund. Denn durch das alternative Außerachtlassen nicht berechenbarer oder ermittelbarer Besteuerungsgrundlagen könnte keine gerechte Besteuerung erreicht werden.“⁵⁰⁶ Demnach münden Mitwirkungsverletzungen, wodurch „die Finanzbehörde die Be-

⁵⁰³ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 54.

⁵⁰⁴ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

⁵⁰⁵ *Webel*, „Steuerfahndung – Steuerstrafverteidigung“, NWB Verlag, 3. Auflage 2016, 124: „Summarische Risikoprüfung und Schnittstellenverprobung: Beide Prüfungsverfahren stellen auf betriebswirtschaftliche und stochastische Gesichtspunkte ab, die zur Verprobung genutzt werden. Ihre Ergebnisse sind aufgrund ihrer umfassenden und systematischen mehrperspektivischen Absicherung ausreichend, um die Ordnungsmäßigkeit der Buchführung zu erschüttern. Darüber hinaus können die Ergebnisse einer Summarischen Risikoprüfung auch durchaus ausreichen, um einen strafprozessualen Anfangsverdacht zu begründen.“ Sowie *Pump/Heidl*, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenbuchführung“, StBp 2014, 162, 163: „Der weitere Vorteil von SRP besteht darin, dass die finanzamtliche Beweisführung für die Hinzuschätzung auch im Steuerstrafverfahren gerichtsfest ist.“

⁵⁰⁶ *Brinkmann*, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017, 21.

steuerungsgrundlagen nicht ermitteln oder berechnen kann⁵⁰⁷, sowie die Aberkennung der „Beweiskraft der Buchführung (und Aufzeichnungen)“⁵⁰⁸ gem. § 162 Abs. 2 AO in der Amtspflicht der Schätzung.

Die Intention einer Schätzung ist es, die Steuern so festzusetzen, als wären die Besteuerungsgrundlagen ordnungsgemäß erklärt worden. Dazu „[steht d]ie Auswahl zwischen verschiedenen Schätzungsmethoden [...] grundsätzlich im pflichtgemäßen Ermessen des FA bzw. FG“⁵⁰⁹. Mit dem erklärten „Ziel, die Besteuerungsgrundlagen durch Wahrscheinlichkeitsüberlegungen so zu bestimmen, dass sie der Wirklichkeit möglichst nahe kommen“⁵¹⁰, hat der BFH eine **Drei-Stufen-Theorie**⁵¹¹ zur Anwendung der gängigsten Schätzungsmethoden aufgestellt. Bei dieser grundsätzlichen Bevorzugung herkömmlicher arbeitsaufwendiger⁵¹² Schätzungsmethoden gegenüber neuen weniger zeitintensiven Ansätzen berücksichtigt der X. Senat jedoch die üblichen **Ermittlungsprobleme der Prüfungspraxis** nicht hinreichend, wonach Geldverkehrs- bzw. Vermögenszuwachsrechnungen sowie Kalkulationen i.d.R. kein genaues Schätzungsergebnis liefern können:

! Für die beiden Formen der **Deckungsrechnung** liegen mangels Aufzeichnungsverpflichtungen keine Informationen zu privaten Geldzuflüssen und -abflüssen vor. „Wenig verständlich ist in diesem Zusammenhang der Hinweis des BFH auf die Möglichkeit des Kontenabrufs. Denn erfahrungsgemäß wird ‚Schwarzgeld‘ in solche Länder verbracht, die an den entsprechenden Verfahren nicht teilnehmen, oder aber es wird mit Hilfe von

⁵⁰⁷ § 162 Abs. 2 S. 1 AO.

⁵⁰⁸ § 158 AO.

⁵⁰⁹ BFH v. 3.9.1998, BFH/NV 1999, 290.

⁵¹⁰ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 60.

⁵¹¹ *Nöcker (RiBFH)*, „Update Betriebsprüfung – Neues vom BFH und die Auswirkungen für die Praxis“, NWB 42/2016, 3157, 3165: „Stufe 1: Die Buchführung ist formell bis auf geringe Mängel korrekt. Folge: Eine Schätzung ist nur bei nachgewiesener materieller Unrichtigkeit der Buchführung möglich. Schätzungsmethoden: Erlösnachkalkulation (Ausbeute- und Aufschlagskalkulation)[,]Geldverkehrsrechnung[und]Vermögenszuwachsrechnung[.]Die aufgedeckten Differenzen können im Rahmen der Schätzung verwendet werden. Stufe 2: Die Buchführung ist formell nicht korrekt, eine materiell fehlerhafte Buchführung ist aber (bislang) nicht konkret nachgewiesen worden. Folge: Eine Schätzung ist aufgrund der folgenden, auf betriebsinternen Daten oder individuellen Verhältnissen des Steuerpflichtigen basierenden Schätzungsmethoden möglich. Schätzungsmethoden: Erlösnachkalkulation (Ausbeute- und Aufschlagskalkulation)[,]Geldverkehrsrechnung[,]Vermögenszuwachsrechnung[und]Zeitreihenvergleich (wenn Steuerpflichtiger nicht mitwirkt oder komplexe finanzielle Verhältnisse vorliegen)[.]Die aufgedeckten Differenzen können im Rahmen der Schätzung verwendet werden, sind jedoch vom Finanzamt – wie auch vom Finanzgericht – von Amts wegen auf Plausibilität anhand der besonderen betrieblichen Verhältnisse des Steuerpflichtigen zu überprüfen, soweit diese bekannt sind. Bei verbleibenden Zweifeln können Sicherheitsabschläge geboten erscheinen. Stufe 3: Die Buchführung ist materiell nicht ordnungsgemäß und folglich auch formell nicht ordnungsgemäß. Folge: Eine Schätzung ist vorzunehmen, die zu einem möglichst genauen Ergebnis kommt, wobei die folgenden Schätzungsmethoden möglich erscheinen: Erlösnachkalkulation (Ausbeute- und Aufschlagskalkulation)[,]Geldverkehrsrechnung[,]Vermögenszuwachsrechnung[und]Zeitreihenvergleich[.]Folge: Die aufgedeckten Differenzen können im Rahmen der Schätzung verwendet werden.“

⁵¹² Vgl. z.B.: BFH v. 21.2.1990, BFH/NV 1990, 683, Rz. 31.: „Die Geldverkehrsrechnung ist eine arbeitsaufwendige Schätzungsmethode.“

Familienangehörigen oder Bekannten bzw. Familienmitgliedern versteckt.“⁵¹³ Darüber hinaus kann mit den Mitteln einer Außenprüfung in den meisten Fällen die nach Prüfungserfahrungen verbreitete Verwendung von un versteuerten Einnahmen für die Bezahlung von ‚Schwarzlöhnen‘ weder ausgeschlossen noch nachgewiesen werden.

! Die für **Kalkulationen** erforderlichen Ausgangsgrößen wie „Wareneinkauf“, „Warenverderb“, „Produktaufschlagsätze bzw. -preise“, „Sonderpreise bzw. -aufschlagsätze“ und „Mengenanteile zu Sonderpreisen“ werden selbst in ordnungsmäßigen Besteuerungsaufzeichnungen nicht vollständig erfasst und finden sich erfahrungsgemäß in den Besteuerungsgrundlagen von Schätzungsfällen nur lückenhaft wieder. Durch die Multiplikation von Fehlern beim Ansatz unrichtiger oder geschätzter Ausgangsgrößen führen Kalkulationen in einer Vielzahl der Fälle zu verstärkten „Hebelwirkungen“⁵¹⁴.

Zudem ist bzgl. einer generellen Festlegung einer Rangfolge von Schätzungsmethoden zu bedenken, dass hierdurch gezielte Unterwanderungsstrategien der systematischen Steuerhinterziehung erleichtert werden.

Brinkmann merkt an: „Bei allem Streben nach ‚Einzelfallgerechtigkeit‘ scheint etwas in den Hintergrund zu geraten, dass die Schätzung ihrer Natur nach gerade keine genaue Berechnung ist und auch nicht sein kann, sondern in erster Linie dazu dient, die Besteuerungsgrundlagen mindestens in einer solchen Höhe festzusetzen, dass kein Vorteil gegenüber demjenigen entsteht, der seinen Verpflichtungen nachgekommen ist.“⁵¹⁵ Dementsprechend wird in der Kommentierung darauf hingewiesen, dass „[n]aturgemäß [...] bei der Schätzung nach Wahrscheinlichkeitsgrundsätzen eine Bandbreite möglicher Wertansätze (sog. **Schätzungsrahmen**) [besteht].“⁵¹⁶

Für die Feststellung des Schätzungsrahmens im Einzelfall und einen sachgerechten Ansatz innerhalb dieses sind die **allgemeinen Maßgaben** bindend: „Der Schätzungsrahmen ist umso größer, je ungesicherter das Tatsachenmaterial ist, auf dem die Schätzung basiert.“⁵¹⁷ Eine aus dem Schätzungsrahmen abgeleitete Schätzung muss vor allem „insgesamt in sich schlüssig [...] wirtschaftlich vernünftig und möglich sein.“⁵¹⁸ Damit das Verifikationsprinzip (s. [Pkt. 4.6.](#)) nicht unterwandert wird, darf „der ‚Beweisverderber‘ oder ‚Beweisvereiteler‘ aus seinem Verhalten keinen Vorteil ziehen [...] (Rechtsgedanke des § 444 ZPO

⁵¹³ *Brinkmann*, „Die Begründung des Sicherheitszuschlags“, StBp 2018, 212, 216.

⁵¹⁴ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 44 ff. zu „Hebelwirkungen“ im Zusammenhang mit Zeitreihenanalysen.

⁵¹⁵ *Brinkmann*, „Die Begründung des Sicherheitszuschlags“, StBp 2018, 212, 218.

⁵¹⁶ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 166 AO, Rz. 44.

⁵¹⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 166 AO, Rz. 44.

⁵¹⁸ BFH v. 18.12.1984, BStBl. II 1986, 226.

i.V.m. § 155 FGO)⁵¹⁹, damit es „keinen Anreiz [gibt], sich weiterhin schätzen zu lassen“⁵²⁰. Folglich „ist das FA im allgemeinen nicht nur berechtigt, sondern verpflichtet, in dem gegebenen Schätzungsrahmen an die oberste Grenze zu gehen.“⁵²¹ „[E]ine überzogene Schätzung [...] hat [...] nur die Rechtswidrigkeit [...] nicht aber bereits ihre Nichtigkeit zur Folge. Etwas anderes ist nach dieser Rechtsprechung, der der Senat folgt, allenfalls zu erwägen, wenn [...] das FA [...] bewusst zum Nachteil des Steuerpflichtigen geschätzt hat.“⁵²²

Die **Quantilsschätzung** ist eine optionale Methode zur Ermittlung der Schätzungshöhe⁵²³ i.R. einer Vollschätzung aber **kein** zwangsweiser Abschluss einer (unaufgeklärt) auffälligen Plausibilitätsprüfung i.S. des § 158 AO. Sie bietet die Möglichkeit, die umfangreichen Erkenntnisse aus der systematisch-mehrperspektivischen SRP für die Feststellung einer sachgerechten Schätzungshöhe zu nutzen. Erfahrungsgemäß ist die informelle Ausgangslage in typischen Fällen, in welchen eine Vollschätzung der Bareinnahmen nach § 162 Abs. 2 AO zu erfolgen hat (s. das *Beispiel unter Pkt. 4.4.*), so lückenhaft bzw. ungesichert, dass auch aufwendige Deckungsrechnungen und Kalkulationen keine genaueren Ergebnisse erzielen können. Zudem bietet die Quantilsschätzung eine verallgemeinerbare Lösung i.S. der Gleichmäßigkeit der Besteuerung gem. § 85 AO, die dennoch variabel auf die Verhältnisse im Einzelfall anpassbar ist.

Weil Steuerhinterziehungen nicht einheitlich sind, sucht die Quantilsschätzung nicht in unrealistischer Erwartung nach einer allgemeingültigen Zeitreihenregel für die wahrscheinlichste Punktschätzung. Sie orientiert sich vielmehr mit ihren **Annahmen und Arbeitsschritten** an den logischen Notwendigkeiten, um aus einer Zeitreihenanalyse mit Hilfe anerkannter Werkzeuge der Mathematik und auf Basis der allgemeinen Maßgaben an eine Schätzung eine sachgerechte Schätzungshöhe abzuleiten:

1. Die Quantilsschätzung setzt auf den betriebseigenen Daten zum Rohgewinnaufschlagsatz und/oder Wareneinsatz auf und ist somit eine **interne Schätzungsmethode**.

⁵¹⁹ BFH v. 15.2.1989, BStBl. II 1989, 462.

⁵²⁰ FG Berlin-Brandenburg v. 30.6.2017 – 10 V 10219/16.

⁵²¹ BFH v. 9.3.1967, BStBl. III 1967, 349; auch gemäß RiBFH Nöcker „zulässig“, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 2015, 3548, 3549.

⁵²² BFH v. 2.12.2000, BStBl. II 2001, 381.

⁵²³ Die Quantilsschätzung ist **keine** Verprobungsmethode, wie es tlw. in der Literatur unzutreffend dargestellt wird; vgl. z.B.: Bleschick, „Der kalkulierte Beanstandungsanlass: Kein Nachweis von Mehrergebnissen durch die Summarische Risikoprüfung – Darstellung der Summarischen Risikoprüfung“, DStR 2017, 353: „Zunächst unternimmt er eine sog. betriebswirtschaftliche Auswertung: Hierzu ermittelt er im Wesentlichen mithilfe des Zeitreihenvergleichs einen Rohgewinnaufschlagsatz. Dieser wird auf den gesamten Wareneinkauf des Steuerpflichtigen angewandt. Unter Außerachtlassung von 20 % der höchsten und ggf. von 20 % der niedrigsten Rohgewinnaufschlagsätze ermittelt er den aus seiner Sicht möglichen Umsatz (sog. „Quantilsschätzung“). Liegt der für den Außenprüfer mögliche Umsatz oberhalb des erklärten Umsatzes, sieht der Außenprüfer hierin aufgrund betriebswirtschaftlicher Auffälligkeiten einen Klärungsbedarf.“

Diese werden den externen allgemein vorgezogen: „Vergleiche mit anderen Betrieben haben regelmäßig einen geringeren Beweiswert als interne Betriebsvergleiche“⁵²⁴. Zusätzlich werden i.R. der Quantilsschätzung gem. § 162 Abs. 1 S. 2 AO „alle Umstände [des Einzelfalls ...] berücksichtig[t]“, die durch Vertiefung oder Mitwirkung konkret bekannt geworden sind und somit in die SRP-Vorlage BWA eingepflegt werden können (s. [Pkt. 4.2.9.](#)).

2. Bei einer zeitreihenbasierten Schätzung muss die zu Grunde zu legende Periode gewählt werden. Diese Entscheidung fällt zwischen den falschen Jahreswerten, welche den Grund für die Vollschätzung bilden, und den empfindlichen Summen kurzer Perioden (s. *das Beispiel zum Einfluss der Zeitreihenperiode auf die Stabilität der Ergebnisse im Abschnitt „Fuzzy-Theorien“ unter [Pkt. 4.2.2.](#)*). Die Quartalsverhältnisse (s. [Pkt. 4.2.4.](#)) sind in Fällen, in denen aufgrund gravierender Aufzeichnungsmängel eine Vollschätzung erforderlich ist, erfahrungsgemäß den falschen Jahreswerten ähnlich, weshalb sie sich für eine sachgerechte Schätzung nicht eignen. **Monatsgrößen** werden in Form der betriebswirtschaftlichen Auswertungen (BWAs) von der BWL, Wirtschaftsprüfung und Steuerberatung als aussagekräftig genug eingeordnet, um existenzrelevante Entscheidungen damit zu begründen wie z.B. Investitionen, Umstrukturierungen, Festanstellungen etc. In der Konsequenz muss ihnen auch die notwendige Robustheit gegenüber alltäglichen⁵²⁵ Störeinflüssen zugesprochen werden, die für eine sachgerechte Schätzung wichtig ist.
3. Auch für relativ robuste Monatskennzahlen kann trotz der gründlichen SRP-Systematik nicht vollständig ausgeschlossen werden, dass einzelne Werte verzerrt sind. Damit diese nicht als Schätzgröße herangezogen werden, müssen sie (möglichst) sicher erkannt und aus dem festzustellenden Schätzungsrahmen⁵²⁶ entfernt werden. **Rangbetrachtungen** sind hierfür deutlich zielführender als arithmetische Verfahren. So würde bspw. ein unrepräsentativ hoher Monatsaufschlagsatz den arithmetischen Mittelwert rechnerisch beeinflussen, nicht aber den mittleren Rang „Median“. Deshalb verwendet die Quantilsschätzung mit dem 80 %-Quantil einen Rang. Derartige relative⁵²⁷ Ränge sind

⁵²⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 162 AO, Rz. 56.

⁵²⁵ *Außergewöhnliche* Störeinflüsse des Einzelfalls werden durch die systematisch-mehrperspektivische Arbeitsweise der SRP aufgedeckt (s. [Pkt. 4.2.2.](#)) und interaktiv auf Normalverhältnisse korrigiert (s. [Pkt. 4.2.9.](#)).

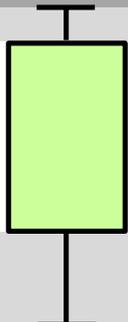
⁵²⁶ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 166 AO, Rz. 44: „Naturgemäß besteht bei der Schätzung nach Wahrscheinlichkeitsgrundsätzen eine Bandbreite möglicher Wertansätze (sog. Schätzungsrahmen). Soweit sich die Schätzung innerhalb des Rahmens bewegt, den der durch die FinBeh. unter Mitwirkung des Stpfl. festgestellte Sachverhalt vorgibt, ist sie nicht zu beanstanden“.

⁵²⁷ Quantile geben die Ränge als Prozentposition in der Datenmenge wieder und werden mit ihrem Namen aus dem Lateinischen abgeleitet: z.B. Quartile (4 Stufen à 25 %), Quintile (5 Stufen à 20 %), Dezile (10 Stufen à 10 %), Perzentile (100 Stufen à 1 %). Fraktile sind Bruchteilsränge, so dass das ¼-Fraktile dem 25 %-Quantil entspricht.

notwendig, wenn evtl. unterschiedlich viele Daten bewertet oder verglichen werden sollen wie bei abweichenden Prüfungszeiträumen.

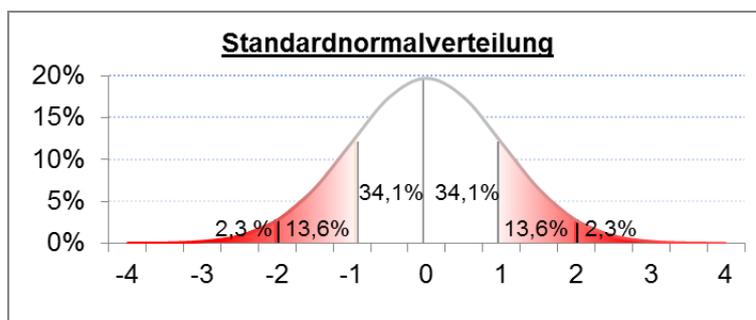
4. Eine übersichtlichere Darstellung der Rangverhältnisse in vorhandenen Aufschlagsätzen bzw. Wareneinsatzsummen liefert das Standardinstrument der Statistik „**Box-Plot**“. Die

nebenstehende Abb. zeigt einen Box-Plot zu zwölf zusammengehörigen Datensätzen. Die Werte werden zunächst (absteigend) sortiert und mit dem absoluten und relativen Rang versehen. Nun werden die Bestandteile „Box“ mit den „normalen“ Daten, „Whisker“ mit den „schwachen Ausreißern“ und darüber hinaus – soweit vorhanden – die „starken Ausreißer“ bestimmt. Diese übliche Form der Quantils-Ordnung von Datenmengen findet sich z.B.

Box-Plot			
Rang	Wert	Prozentrang (Quantil)	Box-Whisker-Plot
12	125	100%	
11	95	91%	
10	55	82%	
9	52	73%	
8	51	64%	
7	49	55%	
6	48	45%	
5	47	36%	
4	46	27%	
3	42	18%	
2	41	9%	
1	38	0%	

auch im Erlass zur Einkunftsabgrenzung zwischen nahestehenden Personen mit grenzüberschreitenden Geschäftsbeziehungen⁵²⁸.

5. Die für eine sachgerechte Schätzung wichtige Schwelle zur Unterscheidung von Normwerten und Ausreißern wird i.R. der Quantilsschätzung aus den **Normalverteilungen** entnommen. Diese beschreiben das normale Streuverhalten natürlich entstandener Daten. Wegen der starken Verbreitung dieser glockenförmigen Verteilungscharakteristik wird sie als „Standardstreuung“ auch dann angenommen, wenn keine anderen Kenntnisse vorliegen. Dabei lässt sich jede Normalverteilung aus den abgebildeten Rang-Wert-Verhältnissen



der Standardnormalverteilung⁵²⁹ ableiten: Den Normalbereich der Daten „Mittelwert ± Standardabweichung“ bilden jeweils ca. 34 % der Datensätze links und rechts vom Mittelwert, also insgesamt ca. 68 %. Die ca. 27 % schwachen Ausreißer stammen aus den Bereichen „-1 bis -2“ und „1 bis 2“ und reichen bis zur doppelten Standardabweichung. Darüber hinaus stellen knapp 5 % starke Ausreißer dar.

⁵²⁸ BMF v. 12.4.2005, BStBl. I 2005, 570, Pkt. 3.4.12.5: „Bandbreiten und ihre Einengung“.

⁵²⁹ Die Standardnormalverteilung ist definiert mit den beiden Kenngrößen „Mittelwert = 0“ und „Standardabweichung (mittlere vorzeichenneutrale Abweichung jedes einzelnen Datensatzes vom Mittelwert) = 1“.

6. In der **Übertragung der Normalverteilungsverhältnisse** auf betriebliche Zahlen fasst die Quantilschätzung den Normalbereich „16 %- bis 84⁵³⁰ %-Quantil“ vorsichtshalber etwas enger mit „20 %- bis 80 %-Quantil“. Dabei ergibt sich für Aufschlagsätze bspw. ein Box-Plot wie in nebenstehender Tabelle: Die niedrigsten und höchsten vier⁵³¹ Aufschlagsätze werden eindeutig als Ausreißer gewertet. Die Ränge „5“ und „20“ liegen tlw. innerhalb und außerhalb der jeweiligen Grenzen „20 %-Quantil“ bzw. „80 %-Quantil“. In solchen Fällen wird zwischen den benachbarten Rängen interpoliert. Die

Quantilsschätzungs-Box-Plot				
RGAS	Monat	Rang	Prozentrang (Quantil)	Box-Whisker-Plot
235%	06/2014	24	100%	
182%	12/2014	23	96%	
178%	07/2015	22	91%	
177%	08/2015	21	87%	
171%	05/2015	20	83%	
169%	03/2015	19	78%	
167%	09/2015	18	74%	
165%	01/2015	17	70%	
156%	12/2015	16	65%	
152%	09/2014	15	61%	
146%	10/2015	14	57%	
145%	11/2015	13	52%	
145%	02/2015	12	48%	
143%	01/2014	11	43%	
142%	04/2015	10	39%	
140%	06/2015	9	35%	
136%	10/2014	8	30%	
131%	03/2014	7	26%	
128%	11/2014	6	22%	
116%	04/2014	5	17%	
115%	08/2014	4	13%	
97%	07/2014	3	9%	
92%	02/2014	2	4%	
88%	05/2014	1	0%	

Quantilsschätzung verwendet hierfür die Standardfunktion von MS Excel[®].⁵³² Dem Verifikationsprinzip⁵³³ folgend wird „in dem gegebenen Schätzungsrahmen [...] die oberste Grenze“⁵³⁴ für die Berechnung der Schätzungshöhe angesetzt.

7. Sofern Informationen vorliegen, dass der Prüfungszeitraum aufgrund gravierender struktureller Veränderungen nicht einheitlich geschätzt werden kann, ist auch eine **geteilte Anwendung der Quantilsschätzung** möglich.⁵³⁵ Im Beispielsfall zeigen die an-

⁵³⁰ Bei der Festlegung des oberen Rands für den Normalbereich müssen in aufsteigender Rangausswertung die unteren Ausreißer mitgezählt werden, also: 16 % untere Ausreißer + 68 % Normalwerte = 84 % obere Grenze des Normalbereichs.

⁵³¹ Bei 24 Monatswerten ergeben sich die beiden Grenzen des Normalbereichs mit: 24 x 20 % bzw. 80 % = 4,8 resp. 19,2.

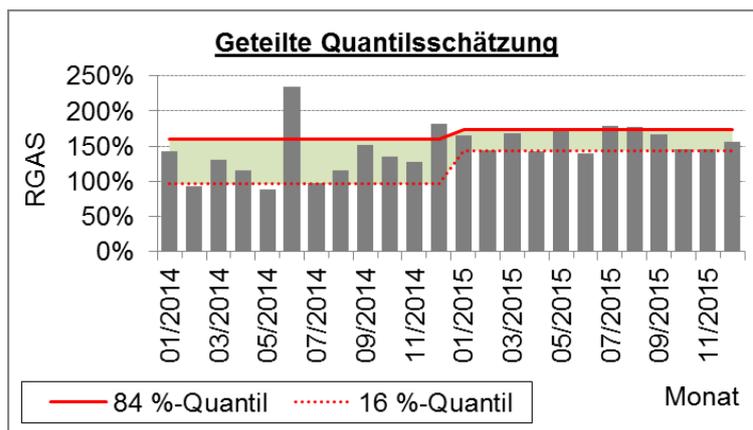
⁵³² Nach den Erfahrungen des Autors mit anderen Tabellenkalkulationsprogrammen wie z.B. OpenOffice[®] führen deren Interpolationsfunktionen zu denselben Ergebnissen.

⁵³³ BFH v. 15.2.1989, BStBl. II 1989, 462: „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO) [...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

⁵³⁴ BFH v. 9.3.1967, BStBl. III 1967, 349; auch gemäß RiBFH Nöcker „zulässig“, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einigemehr“, NWB 2015, 3548, 3549.

⁵³⁵ Die Umsetzung der Quantilsschätzung für uneinheitliche Teilzeiträume lässt sich in der SRP-Vorlage BWA dadurch erreichen, dass zwei (oder mehrere) Vorlagen mit dem kompletten Dateninhalt und jeweils eingeschränktem Prüfungszeitraum verwendet werden. Eine andere Möglichkeit ist die schrittweise Ermittlung der 80 %-Quantile für die Teilzeiträume und die Eingabe im jährlichen Schätzungsfeld „frei geschätzter RGAS“

dersartigen Schwankungen der Aufschlagsätze beider Jahre Unterschiede im Geschäftsbetrieb an. Deshalb werden für beide Prüfungsjahre getrennte Normalbereiche ermittelt. Durch die Verwendung von relativen Rängen erfolgt mit dem



80 %-Quantil „19,2-ter größter Wert“ bei 24 Monaten ein vergleichbar hoher Ansatz wie mit dem 9,6-ten bei 12 Monaten.

- Abschließend ist das Schätzungsergebnis überschlägig daraufhin zu prüfen, ob es „insgesamt in sich schlüssig [...] wirtschaftlich vernünftig und möglich“⁵³⁶ ist. Diese **Gegenkontrolle** kann z.B. durch eine Einschätzung des Preisgefüges erfolgen.

Folgende Hinweise zur Durchführung einer Quantilsschätzung mit der SRP-Vorlage BWA sind wichtig:

➔ Die **Ermittlung der** **80 %-Quantil der Monatswerte:** **302%** **8.079,10**

Schätzgrößen „80 %-Quantil“ (s. den Aufzählungspkt. 6 „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.)

zu den monatlichen Aufschlagsätzen und Wareneinkaufs- bzw. -einsatzsummen kann durch **Aktivierung** in einem Extrateil des Tabellenblatts „Monatszeitreihen“ dargestellt werden. Zusätzlich zur Anzeige der beiden Werte erscheint neben dem Diagramm der Quantilsschätzungs-Box-Plot mit dem Normalbereich und den **formatseitig hervorgehobenen eindeutigen Ausreißern**. Durch **Ankreuzen** kann zwischen den Monatsgrößen für die Aufschlagsätze und für den Wareneinkauf bzw. -einsatz gewählt werden. Die zu Grunde liegenden

Ranking RGAS	
Rang	Wert
36	433 %
35	391 %
34	388 %
33	376 %
32	355 %
31	337 %
30	302 %
29	302 %
28	267 %
27	257 %
26	253 %
25	249 %
24	245 %
23	244 %

Betriebszahlen beinhalten bei Aktivierung die gleichmäßig umverteilten Warenentnahmen vom Tabellenblatt „Jahressummen“ (s. [Pkt. 4.2.3.](#))⁵³⁷ Zudem werden alle interakti-

bzw. „frei geschätzter Einsatz“ (s. den Auflistungspkt. „freie Schätzung“ im folgenden Abschnitt „Hinweise zur Durchführung einer Quantilsschätzung“).

⁵³⁶ BFH v. 18.12.1984, BStBl. II 1986, 226.

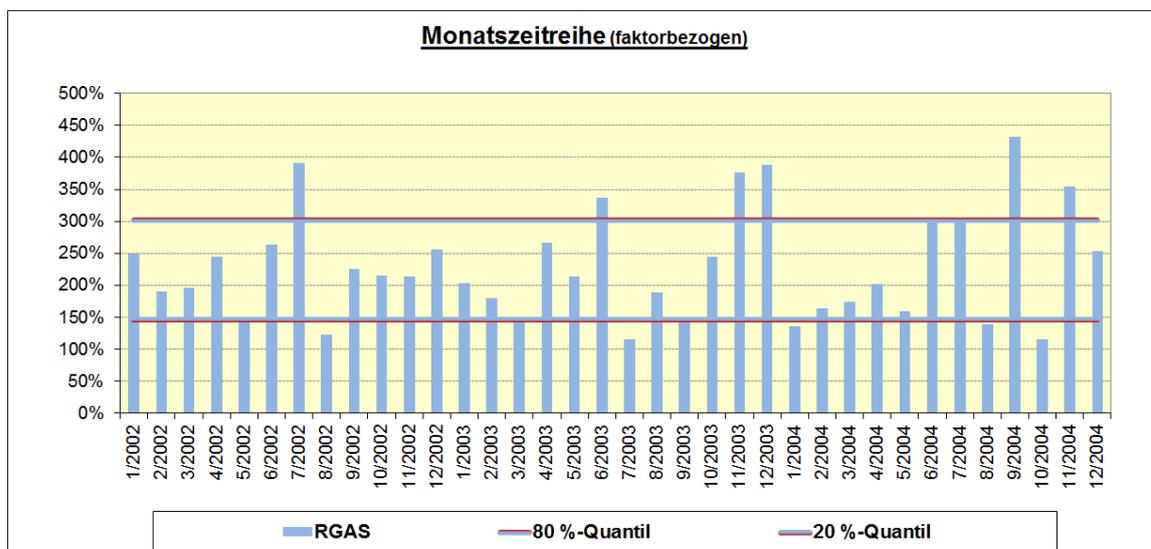
⁵³⁷ Eine entsprechende (gleichmäßige) Umverteilung der auf dem Tabellenblatt „Jahressummen“ eingetragenen jährlichen Bestandsveränderungen ist in der SRP-Vorlage BWA nicht vorgesehen, weil dies in aller Regel nicht den tatsächlichen Abläufen entspricht. Sollen monatliche Bestandsveränderungen in die Ermittlung der

ven Korrekturen (s. [Pkt. 4.2.9.](#)) wie insbesondere Umverteilungen oder Berichtigungen von Einzeldatensätzen (s. [Pkt. 4.2.9.1.](#)), eine aktivierte Kompletterschiebung der ersten Datenreihe (s. [Pkt. 4.2.9.3.](#)) und die Konjunkturanpassungen auf dem Tabellenblatt „Monatszeitreihen“ (s. [Pkt. 4.2.5.1.](#) u. [Pkt. 4.2.9.4.](#)) berücksichtigt.⁵³⁸ Durch Warnhinweise am zentralen Schätzungsfeld „80 %-Quantils RGAS:“ auf dem Tabellenblatt „Schätzung“ wird zur Absicherung an die vorgenommenen Eingriffe erinnert.

Sollen einzelne Monate bei der Ermittlung der 80 %-Quantile ausgeschlossen	2005					
	Monat	Umsatz		Einkauf		RGAS
	Jan	27.308,55	4%	2.488,49	2%	<input checked="" type="checkbox"/> 997%
	Feb	28.553,28	5%	8.306,82	5%	<input type="checkbox"/> 244%
	Mär	40.280,83	7%	7.602,51	5%	<input type="checkbox"/> 430%
	Apr	49.500,40	8%	16.904,47	10%	<input type="checkbox"/> 193%
	Mai	66.381,87	11%	13.902,28	8%	<input type="checkbox"/> 377%

werden, insbesondere weil es sich bei Saisongeschäften um Vorbereitungszeiten ohne vergleichbaren operativen Betrieb handelt, können sie [nach Aktivierung einzeln abgewählt](#) werden.

Über die Angabe der Werte und die Tabelle hinaus wird der Schätzungsrahmen in dem Extrabereich des Tabellenblatts „Monatszeitreihen“ als Grafik ausgegeben:



Die Visualisierung des festgestellten Schätzungsrahmens und der abzugrenzenden Ausreißer dient dem Zweck, das Schätzungsergebnis übersichtlich zu kontrollieren (s. *den Aufzählungspkt. 8 „Gegenkontrolle“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.*). Hierbei kann besonders die Notwendigkeit zu getrennten Quantilsschätzungen für Teilzeiträume erkannt bzw. ausgeschlossen werden (s. *den Aufzählungs-*

80 %-Quantile für die Aufschlagsätze bzw. den Wareneinkauf bzw. -einsatz einbezogen werden, müssen sie – anstatt auf dem Tabellenblatt „Jahressummen“ – im Korrekturbereich des Tabellenblatts „Vergleichsdaten“ eingetragen werden.

⁵³⁸ Die Funktion „Niveaueinpassung eines Grafen“ im Bereich „einzelskalierter Zeitreihenvergleich“ auf dem Tabellenblatt „Monatszeitreihen“ (s. [Pkt. 4.2.5.2.](#) u. [Pkt. 4.2.9.5.](#)) hat keinen Einfluss auf die Monatszahlen, die zur Ermittlung der Schätzgrößen „80 %-Quantil“ verwendet werden.

pkt. 7 „geteilte Anwendung der Quantilsschätzung“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.).

- ➔ Das 80 %-Quantil der Aufschlagsätze kann auch im oberen Diagramm der **Gleitschlitzenzeitreihen** (s. [Pkt. 4.2.6.](#)) angezeigt werden. Unabhängig von der Auswahl der Gleitschlitzenperiode wird hierbei immer das 80 %-Quantil der Monatsaufschlagsätze angezeigt, weil diese nach logischen Erwägungen die Grundlage der Quantilsschätzung bilden (s. den Aufzählungspkt. 2 „Monatsgrößen“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.). Auch diese Verbildlichungsfunktion soll zur Absicherung beitragen, ob der gefundene Aufschlagsatz für eine Schätzung „insgesamt in sich schlüssig [...] wirtschaftlich vernünftig und möglich“⁵³⁹ ist.
- ➔ Die eigentliche Quantilsschätzung findet auf dem **Tabellenblatt „Schätzung“** statt:

Interaktive Schätzungsvorlage				
Steuerpflichtige/r:	Restaurant		Steuernummer:	12 345 6789 0
	2002	2003	2004	
Umsatz:	273.498,86	276.713,18	234.729,54	
Einkauf:	87.895,17	84.665,18	76.496,22	
Bestandsveränderung:	4.904,47	-5.575,55	2,75	
Entnahmen:	2.460,00	2.460,00	2.592,00	
zusätzliche Entnahmen (ermäßigter USt-Satz):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
zusätzliche Entnahmen (voller USt-Satz):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
weitere Einsatzminderungen:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Einsatz:	80.530,70	87.780,73	73.901,47	
RGAS:	240%	215%	218%	
80 %-Quantil RGAS:	<input type="text"/>	<input type="text"/>	<input type="text"/>	
frei geschätzter RGAS:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
80 %-Quantil Einkauf / Einsatz:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
frei geschätzter Einsatz:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
ermittelter Umsatz:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
freie Umsatzhinzuschätzung:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Umsatzschätzung:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Mehrerlös durch Schätzung:	<input type="checkbox"/>			
(RGAS nach Schätzung):	<input type="checkbox"/>			
Mehrerlös insgesamt:	<input type="checkbox"/>			
Anteil ermäßigter Steuersatz:	<input type="text"/>	<input type="text"/>	<input type="text"/>	
Anteil Mehrerlös (ermäßigter USt-Satz):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
freie Aufteilung Mehrerlös:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Anteil Mehrerlös (ermäßigter USt-Satz):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Anteil Mehrerlös (voller USt-Satz):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Mehr-USt durch Schätzung:	<input type="checkbox"/>			
Mehr-USt insgesamt:	<input type="checkbox"/>			

Im hellblauen Bereich werden die Jahressummen inkl. aller vorgenommenen Korrekturen (s. den Auflistungspkt. „Ermittlung der Schätzgrößen ,80 %-Quantil“ oberhalb) übernommen. Der gesamte rot hinterlegte Bereich ist ausschließlich für die (gemeinsame) Erarbeitung einer sachgerechten Schätzungshöhe vorgesehen. Sofern eine Einigung über die Hinzuschätzung erreicht werden kann, lässt sich das Vorlagenblatt über das oberste Ankreuzfeld in „Anlage zur Tatsächlichen Verständigung“ umbenennen. Auf diese Weise kann der letzte gemeinsame Verhandlungsstand unproblematisch dokumentiert und der Verfahrensablauf nach einer Tatsächlichen Verständigung somit erleichtert werden.

- ➔ Wie in der Abb. zum Auflistungspkt. „Tabellenblatt ,Schätzung“ zu erkennen, hat das zentrale Schätzungsfeld „80 %-Quantil RGAS:“ eine **Warnfunktion:**

⁵³⁹ BFH v. 18.12.1984, BStBl. II 1986, 226.

Durch eine **rote Umrandung des Feldes** wird darauf hingewiesen,

1. dass entweder auf dem Tabellenblatt „Jahressummen“ eingetragene Entnahmen nicht per Aktivierung auf dem Tabellenblatt „Monatszeitreihen“ umverteilt worden sind (s. [Pkt. 4.2.3.](#)),
2. sich Einzeldaten-Korrekturen auf dem Tabellenblatt „Vergleichsdaten“ in ihrer Summe nicht als reine Umverteilung ausgleichen (s. [Pkt. 4.2.9.1.](#)) oder
3. einzelne Datensätze außerhalb des auf dem Tabellenblatt „Vergleichsdaten“ eingegebenen Prüfungszeitraums liegen und somit nicht berücksichtigt werden können.

Die **Rotverfärbung der Schrift** „80 %-Quantil RGAS:“ erinnert daran,

1. dass entweder die erste Datenreihe auf dem Tabellenblatt „Vergleichsdaten“ komplett vor- oder rückdatiert wird (s. [Pkt. 4.2.9.3.](#)) oder
2. die untersuchten Daten mittels Konjunkturanpassung auf dem Tabellenblatt „Monatszeitreihen“ (tlw.) von einem nominalen auf ein reales Preisniveau umgerechnet werden (s. [Pkt. 4.2.5.1.](#) u. [Pkt. 4.2.9.4.](#)).

Die Quantilsschätzung wird jedoch infolge der erkannten Eingriffe in die Betriebsdaten nicht unterbunden, weil gerade das Beachten „alle[r] Umstände [...], die für die Schätzung von Bedeutung sind“ gem. § 162 Abs. 1 S. 2 AO zu einem sachgerechten Schätzungsergebnis wesentlich beiträgt. Vielmehr soll die Erinnerung nach einem umfangreichen Prüfungsprozess verhindern, dass Datenveränderungen **ungewollt** die Schätzung beeinflussen.

➔ *Im Gegensatz zu Kalkulationen hätte eine Erhöhung der Entnahmen bei der Schätzung mit einem bestimmten Aufschlagsatz-Quantil keine sichere ergebnismindernde Wirkung, weil die Verringerung des umsatz erzeugenden Wareneinsatzes zu durchgehend höheren Aufschlagsätzen führt und damit i.d.R. auch das betreffende z.B. 80 %-Quantil vergrößert wird.*

		2002
Umsatz:		273.498,86
Einkauf:		87.895,17
Bestandsveränderung:		4.904,47
Entnahmen:		2.460,00
zusätzliche Entnahmen (ermäßigter USt-Satz):	<input checked="" type="checkbox"/> 7%	1.000,00
zusätzliche Entnahmen (voller USt-Satz):	<input checked="" type="checkbox"/> 16%	1.500,00
weitere Einsatzminderungen:		
Einsatz:		78.030,70
RGAS:		251%
80 %-Quantil RGAS:	<input checked="" type="checkbox"/>	317%

Um die Möglichkeit einvernehmlicher Schätzungen nicht zu verschlechtern, bietet die SRP-Vorlage BWA auf dem Tabellenblatt „Schätzung“ eine zielführende Funktion an, **zusätzliche Entnahmen im Schätzungsweg** zu ergänzen. Eingaben, welche im

Bereich „zusätzliche Entnahmen“ oder „weitere Einsatzminderungen“ eingetragen werden, haben aus verhandlungstechnischen Gründen keine Auswirkungen auf die Ermittlung der **80 %-Quantile** (s. den Auflistungspkt. „Ermittlung der Schätzgrößen ,80 %-Quantil“ oberhalb). Für die Berechnung der jeweiligen Schätzungsauswirkungen (s. den Auflistungspkt. „Berechnung des Schätzungsergebnisses“ unterhalb) können die Werte durch **Aktivierung** in die Umsatzbesteuerung einbezogen werden; der **Umsatzsteuersatz** wird dazu von der SRP-Vorlage BWA automatisch dem (Wirtschafts-)Jahr⁵⁴⁰ zugesteuert.

- ➔ Soll (auch) der **Wareneinsatz mit dem 80 %-Quantil** geschätzt werden⁵⁴¹, weil eindeutige Hinweise auf eine Doppelverkürzung vorliegen, kann dies ebenfalls per Anhaken auf dem Tabellenblatt „Schätzung“ der SRP-Vorlage BWA geschehen:

80 %-Quantil RGAS:	<input type="checkbox"/>	
frei geschätzter RGAS:	<input type="checkbox"/>	
80 %-Quantil Einkauf / Einsatz:	<input checked="" type="checkbox"/>	12 94.489,20

Es wird automatisch die **Anzahl an Monaten** mit einem Wareneinkauf bzw. -einsatz, der größer als Null ist, übernommen. Dieser Multiplikator für den 80 %-höchsten Monatswareneinkauf bzw. -einsatz kann manuell abgeändert werden, wenn in einzelnen Monaten kein oder nur ein deutlich reduzierter Geschäftsbetrieb z.B. aufgrund von Urlaub stattgefunden hat.

- ➔ Die SRP-Vorlage BWA unterstützt mit ihrer transparenten Darstellung einer Schätzungsbesprechung ebenfalls die **freie Schätzung** des Aufschlagsatzes und/oder Wareneinsatzes. Dazu können die Jahreswerte einzeln eingegeben oder per Anhakfeld vom ersten Jahr auf alle übertragen werden. Sobald widersprüchliche Eingaben vorliegen wie bspw. die aktivierte Quantilsschätzung für den Wareneinsatz und ein manuell eingetragener Wert, unterbindet die SRP-Vorlage BWA den Auswertungsprozess mit einem entsprechenden Fehlerhinweis. Das abschließende Feld im Auswahlblock der Erlösschätzung „**freie Umsatzhinzuschätzung**“ übergeht in der Ergebnisberechnung (s. den Auflistungspkt. „Berechnung des Schätzungsergebnisses“ unterhalb) alle anderen Eingaben, weil dieser Schritt in Schätzungsverhandlungen in aller Regel den letzten (einvernehmlichen) Stand darstellt, der zu dokumentieren ist.

80 %-Quantil RGAS:
frei geschätzter RGAS:
80 %-Quantil Einkauf / Einsatz:
frei geschätzter Einsatz:
ermittelter Umsatz:
freie Umsatzhinzuschätzung:

⁵⁴⁰ Für das Wirtschaftsjahr 2007 wird der volle Umsatzsteuersatz von der SRP-Vorlage BWA tageanteilig ermittelt, kann aber manuell abgeändert werden.

⁵⁴¹ Hierzu wird der 80 %-höchste Rang der monatlichen Wareneinkaufs- bzw. -einsatzsummen ermittelt (s. den Auflistungspkt. „Ermittlung der Schätzgrößen ,80 %-Quantil“ oberhalb) und mit der Anzahl an Monaten multipliziert.

- ➔ Sobald auswertungsfähige Daten in die SRP-Vorlage BWA hineinkopiert worden sind, ein Prüfungszeitraum eingegeben und eine Schätzungsform ausgewählt wurden, erfolgt die **Berechnung des Schätzungsergebnisses** pro Jahr und insgesamt:

80 %-Quantil RGAS:	<input type="checkbox"/>				
frei geschätzter RGAS:	<input checked="" type="checkbox"/>				
80 %-Quantil Einkauf / Einsatz:	<input type="checkbox"/>				
frei geschätzter Einsatz:	<input type="checkbox"/>				
ermittelter Umsatz:	<input type="checkbox"/>			289.676,41	243.874,85
freie Umsatzhinzuschätzung:	<input type="checkbox"/>				
Umsatzschätzung:	<input type="checkbox"/>			289.676,41	243.874,85
Mehrerlös durch Schätzung:	<input type="checkbox"/>			12.963,23	9.145,31
(RGAS nach Schätzung:)	<input type="checkbox"/>				
Mehrerlös insgesamt:	<input type="checkbox"/>				14.360,99
Anteil ermäßigter Steuersatz:	<input type="checkbox"/>			30%	
Anteil Mehrerlös (ermäßigter USt-Satz):	<input type="checkbox"/>	7%	-2.324,26	7%	2.743,59
freie Aufteilung Mehrerlös:	<input type="checkbox"/>				
Anteil Mehrerlös (ermäßigter USt-Satz):	<input type="checkbox"/>				
Anteil Mehrerlös (voller USt-Satz):	<input type="checkbox"/>	16%	-5.423,28	16%	6.401,72
Mehr-USt durch Schätzung:	<input type="checkbox"/>			1.724,11	1.216,33
Mehr-USt insgesamt:	<input type="checkbox"/>				1.910,01

Kommt es durch die Schätzungseingaben zu **Minderschätzungen** wird formatseitig durch rote Schrift darauf hingewiesen. Der Berechnungsprozess wird allerdings nicht abgebrochen, weil ein partielles Negativergebnis zielführender Bestandteil einer Gesamtschätzung sein kann.

Im unteren Abschnitt des Tabellenblatts „Schätzung“ kann neben der Ermittlung der Umsatzhinzuschätzung die Auswirkung auf die Umsatzsteuer berechnet werden. Dazu wird der Mehrerlös entweder **einheitlich per Prozentsatz oder durch jahresweise Eingabe des Betrags zum ermäßigten Umsatzsteuersatz** aufgeteilt. Wareneinsatzminderungen, die i.R. der Schätzung ergänzt worden sind (*s. den Auflistungspkt. „zusätzliche Entnahmen im Schätzungsweg“ oberhalb*), werden durch Anhaken dabei berücksichtigt.

Nach der länderübergreifenden Verbreitung der Quantilsschätzung haben einige Finanzgerichte (z.T. mehrfach) zu dem Verfahren Stellung genommen.⁵⁴² Dabei wird es tlw. ausführlich und korrekt wiedergegeben.⁵⁴³ In den überwiegenden Fällen wurde die Quantilsschätzung von der **Finanzrechtsprechung** gegenüber den in Vollschätzungsfällen zulässigen griffweisen Schätzungen⁵⁴⁴ als „deutlich differenziertere Schätzungsmethode“⁵⁴⁵ anerkannt. Insbesondere der BFH⁵⁴⁶ und der 5. Senat des FG Berlin-Brandenburg⁵⁴⁷ haben ihre skepti-

⁵⁴² Dem Autor bekannt sind: FG Hamburg v. 5.3.2018 – 3 K 205/15, FG Hamburg v. 18.7.2017 – 6 V 119/17, FG Berlin-Brandenburg v. 30.6.2017 – 10 V 10219/16, FG Berlin-Brandenburg v. 9.1.2017 – 4 V 4255 u. 4265/15, FG Hamburg v. 31.10.2016 – 2 V 202/16, FG Hamburg v. 26.8.2016 – 6 V 81/16, FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16, FG Berlin-Brandenburg v. 25.5.2015 – 1 V 1039/15, FG Schleswig-Holstein v. 30.9.2014 – 1 V 83/14, FG Berlin-Brandenburg v. 31.1.2013 – 14 V 14316/11 u. FG Sachsen v. 10.1.2012 – 6 V 269/11.

⁵⁴³ Vgl. z.B.: FG Hamburg v. 26.8.2016 – 6 V 81/16, Rz. 85 ff.

⁵⁴⁴ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16, 4. Leitsatz.

⁵⁴⁵ FG Berlin-Brandenburg v. 30.6.2017 – 10 V 10219/16.

⁵⁴⁶ Beschl. v. 12.7.2017 – X B 16/17; da nach Auffassung des X. Senats gem. Rz. 86 im betreffenden Fall „bei summarischer Betrachtung keine Befugnis zur Vollschätzung abgeleitet werden [kann]“, war die Stellungnahme zur durchgeführten Quantilsschätzung als reine Schätzungsmethode nicht erforderlich.

⁵⁴⁷ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16.

sche Haltung gegenüber der Quantilsschätzung auf die folgenden **Kritikpunkte** gestützt, die jedoch – wie dargestellt – inhaltlich nicht überzeugend sind:

- ! „Die besonderen Risiken [einer Zeitreihenschätzung] ([...], erhebliche mathematische Hebelwirkungen, [...]) bestehen hier ebenso.“⁵⁴⁸
 - ➔ Die kritischen Ausführungen des BFH zu möglichen „Hebelwirkungen“ infolge von Zuordnungsfehlern in Zeitreihenanalysen stammen aus punktuellen Betrachtungen des maximalen 10-Wochen-Aufschlagsatzes.⁵⁴⁹ I.R. der Quantilsschätzung werden die stöempfindlichen Extremwerte sorgfältig ausgeschlossen (s. die *Aufzählungspkte. 4-6 „Box-Plot“, „Normalverteilungen“ und „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.*). In der BWL, Wirtschaftsprüfung und Steuerberatung werden normale⁵⁵⁰ Monatskennzahlen eines Unternehmens als so robust gegenüber Alltagseinflüssen angesehen, dass sie die Grundlage für existentielle Entscheidungen bilden (s. *den Aufzählungspkt. 2 „Monatsgrößen“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.*). Evtl. verzerrende Sondersachverhalte werden vor einer Quantilsschätzung mit Hilfe der SRP-Prüfungssystematik gezielt gesucht und korrigiert (s. [Pkt. 4.2.2.](#)).
- ! „Die besonderen Risiken [einer Zeitreihenschätzung] ([...], Ausgabe großer und kaum vollständig überprüfbarer Datenmengen) bestehen hier ebenso.“⁵⁵¹
 - ➔ Im Gegensatz zu den kritisierten „ca. 1 100 Zeilen zu je 10 Spalten, d.h. insgesamt ca. 11 000 Eintragungen“⁵⁵² setzt die Quantilsschätzung i.d.R. auf 36 Monatswerten auf. Diese Zahlen resultieren aus den zur Prüfung überlassenen Daten und möglichen interaktiven Korrekturen (s. [Pkt. 4.2.9.](#)), welche durch die SRP-Vorlagen nachvollziehbar dokumentiert werden und z.T. das Ergebnis aus der Mitwirkung Stpfl. sind. Zudem ist die verringerte Transparenz bei der Prüfung von digitalen Betriebsdaten vielmehr eine logische Konsequenz daraus, dass „die Anzahl elektronisch überwachter Prozesse stetig [steigt].“⁵⁵³
- ! „Das FA hat hierzu in der Beschwerdeerwiderung ausgeführt, mit der ‚Konjunkturanpassung‘ hätten die Ergebnisse der Jahre 2009 und 2010 an das Preisniveau des Jahres 2008 angepasst werden sollen, um die Vergleichbarkeit der Schätzung herzustellen. In- des soll eine Schätzung nicht vergleichbar, sondern möglichst realitätsnah sein.“⁵⁵⁴

⁵⁴⁸ BFH v. 12.7.2017 – X B 16/17, Rz. 111.

⁵⁴⁹ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 44 ff.

⁵⁵⁰ Nach Ausschluss von offensichtlichen unrepräsentativen Ausreißern.

⁵⁵¹ BFH v. 12.7.2017 – X B 16/17, Rz. 111.

⁵⁵² BFH v. 25.3.2015, BStBl. II 2015, 743, Rz. 50.

⁵⁵³ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876 u. 877.

⁵⁵⁴ BFH v. 12.7.2017 – X B 16/17, Rz. 128.

- ➔ Die Bereinigung von Preisveränderungen mittels Preisindizes ist das übliche Werkzeug der Volkswirtschaftslehre, um Zahlen über die Jahre vergleichen zu können. Auch i.R. der SRP dient dieses Mittel dazu, besonders einseitige Preiseffekte zu korrigieren (s. [Pkt. 4.2.5.1.](#) u. [Pkt. 4.2.9.4.](#)), damit die Zeitreihenanalysen hierdurch nicht verzerrt werden. Im betreffenden Fall hat der Prüfer die Konjunkturanpassung in der Gesamtbildprüfung verwendet, um die Schlüssigkeit und Abhängigkeit des Warenflusses unabhängig von der Preiserhöhung der Gaststätte gem. § 158 AO beurteilen zu können. Für die Quantilsschätzung hat er die Konjunkturbereinigung wieder aufgelöst und somit die tatsächlichen Verhältnisse zu Grunde gelegt.⁵⁵⁵
- ! „Dieser Einwand verkennt, dass die Methode der Quantilsschätzung gerade auf der Heranziehung eines der höchsten RAS beruht. [...] Tatsächlich rekurriert die Quantilsschätzung nicht etwa auf den ‚normalen Geschäftsverlauf‘, sondern stützt sich für die vorgenommene Vollschatzung auf einen Wert, der in 80 % der Zeitabschnitte gerade nicht erreicht wird.“⁵⁵⁶
- ➔ Innerhalb des anhand von Betriebsdaten feststellbaren Schätzungsrahmens (s. *den Abschnitt „Schätzungsrahmen“ weiter oben unter diesem Pkt.*) wählt die Quantilsschätzung *vorsichtig* die „oberste Grenze“⁵⁵⁷ (s. *die Aufzählungspkt. 6 „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.*), weil gem. Verifikationsprinzip (s. [Pkt. 4.6.](#)) „der ‚Beweisverderber‘ oder ‚Beweisvereiteler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO)⁵⁵⁸. Ein mittlerer Ansatz im Schätzungsrahmen würde nicht zu einem wahrscheinlicheren Schätzungsergebnis führen⁵⁵⁹, sondern dem Ziel entgegenwirken, dass es „keinen Anreiz [gibt], sich weiterhin schätzen zu lassen“⁵⁶⁰.
- ! „In mathematischer Hinsicht setzt die Anwendung [...] zuvörderst voraus, dass die RAS überhaupt der Normalverteilung folgen und die erhobene Grundgesamtheit (hier: 36 Einzelgrößen) groß genug ist.“⁵⁶¹
- ➔ Die Ableitung eines sachgerechten Schätzungsergebnisses aus den betriebseigenen Zeitreihengrößen setzt den (möglichst sicheren) Ausschluss von verzerrten

⁵⁵⁵ Vgl. BFH v. 12.7.2017 – X B 16/17, Rz. 10 ff.

⁵⁵⁶ BFH v. 12.7.2017 – X B 16/17, Rz. 135 u. 144.

⁵⁵⁷ BFH v. 9.3.1967, BStBl. III 1967, 349; auch gemäß RiBFH Nöcker „zulässig“, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 2015, 3548, 3549.

⁵⁵⁸ BFH v. 15.2.1989, BStBl. II 1989, 462.

⁵⁵⁹ Durch einen mittleren Ansatz im Schätzungsrahmen würde das Ergebnis den erklärten Besteuerungsgrundlagen nahekommen und sie somit nachträglich anerkennen, obwohl sie nach der Gesamtbildprüfung gem. § 158 AO zu verwerfen waren.

⁵⁶⁰ FG Berlin-Brandenburg v. 30.6.2017 – 10 V 10219/16.

⁵⁶¹ BFH v. 12.7.2017 – X B 16/17, Rz. 139.

Ausreißern voraus. Solange keine anderen Erkenntnisse zum Streuverhalten von Daten bekannt sind, wird zunächst regelmäßig das bekannteste der Normalverteilungen angenommen (s. die Aufzählungspkt. 5 „Normalverteilungen“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.). Ob Rohgewinnaufschlagsätze grundsätzlich normalverteilt sind, ist dafür unerheblich, weil die verschiedenen Hinterziehungsformen unterschiedliche Auswirkungen auf die Verteilung hätten. Insofern ist der Ansatz der Normalverteilungen logisch. Da im Quantilsschätzungs-Box-Plot mit Quantilen gearbeitet wird (s. die Aufzählungspkt. 6 „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.), ist zudem die Anzahl an vorliegenden Monatswerten unbeachtlich, denn das Ziel von relativen Prozenträngen ist gerade die Übertragbarkeit auf unterschiedliche und auch wenige Daten.

- ! „Das FG Hamburg (Beschluss vom 26. August 2016 6 V 81/16) hatte in einem AdV-Beschluss einen Sachverhalt zu beurteilen, in dem Schwarzeinkäufe konkret nachgewiesen waren, die Gewinnermittlung sich also schon ohne den Zeitreihenvergleich als materiell unrichtig erwiesen hatte. Es hat hier die Quantilsschätzung dem Grunde nach nicht beanstandet, der Höhe nach aber erhebliche Korrekturen vorgenommen.“⁵⁶²
 - ➔ Tatsächlich wurden die Hinzuschätzungen durch die Quantilsschätzung mit einem Aufschlagsatz von 199 % im betreffenden Beschluss vollständig anerkannt.⁵⁶³ Die herausgestellten „erheblichen Korrekturen“ beziehen sich auf andere Sachverhalte.
- ! „Insoweit ist *Kulosa*⁵⁶⁴ zuzustimmen, wenn dieser ausführt, dass ‚rechnerische oder grafische Ergebnisse selbst noch keine zwingenden Besteuerungsgrundlagen darstellen, sondern zunächst in den tatsächlichen Lebenssachverhalt eingeordnet, auf Plausibilität verprobt und rechtlich gewürdigt werden müssen‘.“⁵⁶⁵
 - ➔ Diese Generalkritik richtet sich gleichermaßen gegen alle indirekten Prüfungsverfahren, also auch gegen die höchstrichterlich anerkannten Vermögenszuwachs-, Geldverkehrsrechnungen und Kalkulationen⁵⁶⁶. Da das Verifikationsprinzip⁵⁶⁷ gera-

⁵⁶² BFH v. 12.7.2017 – X B 16/17, Rz. 115.

⁵⁶³ FG Hamburg v. 26.8.2016 – 6 V 81/16, Rz. 19 u. 97.

⁵⁶⁴ *Kulosa (RiBFH)*, „Mathematisch-statistische Schätzungsmethoden in der Betriebsprüfung – Erste BFH-Rechtsprechung zum Zeitreihenvergleich“, DB 2015, 1797, 1799.

⁵⁶⁵ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16.

⁵⁶⁶ BFH v. 25.3.2015, BStBl. II 2015, 743, Rz.

⁵⁶⁷ BFH v. 16.5.2013, BStBl. II 2014, 225: „Der Gesetzgeber ist demgemäß von Verfassungs wegen verpflichtet, zur Vermeidung der Verfassungswidrigkeit des materiellen Steuergesetzes dieses in ein normatives Umfeld einzubetten, das die tatsächliche Lastengleichheit der Steuerpflichtigen gewährleistet, insbesondere auch durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip.“

de die Anwendung von verprobenden⁵⁶⁸ Gesamtbildmethoden⁵⁶⁹ vorsieht, um nicht nur im Einzelfall sondern auch in der Gesamtheit der Stpfl. die Gesetzmäßigkeit der Besteuerung nach § 85 AO umzusetzen, geht die Kritik ins Leere.

- ! „Die Quantilsermittlung stellt nach Auffassung des Senats lediglich eine Verprobungs-, aber keine zulässige Schätzungsmethode dar. Denn ebenso wie andere Verprobungsmethoden [...] zeigt die Quantilsermittlung lediglich sog. Ausreißerwerte an, die sodann zu weiteren Ermittlungen Anlass geben, [...]“⁵⁷⁰.
- ➔ Von Beginn an⁵⁷¹ ist die Quantilsschätzung auch begrifflich mit dem „objektivierten Leistungsfähigkeitsmaß“⁵⁷² als reine Schätzungsmethode vorgestellt worden. Dabei hat das (möglichst vollständige) Erkennen von unrepräsentativen Ausreißerwerten in den der Schätzung zu Grunde zu legenden Betriebsdaten eine überragende Bedeutung dafür, dass das berechnete Schätzungsergebnis nicht verzerrt sondern sachgerecht ist (s. die Aufzählungspkte. 4-6 „Box-Plot“, „Normalverteilungen“ und „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.).
- ! „Vielmehr stellt es nach Auffassung des Senats einen Fehlschluss dar, dass zwischen der Eliminierung von Ausreißerwerten sowie der sich daraus ergebenden Ermittlung eines Normalmittels einerseits und den tatsächlichen betrieblichen Gegebenheiten andererseits ein sachlicher Zusammenhang besteht, [...]“⁵⁷³.
- ➔ Der „[Zusammenhang] zwischen der Eliminierung von [unrepräsentativen] Ausreißerwerten [...] und den tatsächlichen betrieblichen Gegebenheiten“ besteht darin, dass der aus den Betriebsdaten verbleibende Schätzungsrahmen (s. den Abschnitt „Schätzungsrahmen“ weiter oben unter diesem Pkt.) als vernünftiger Ausgangspunkt für eine Schätzung⁵⁷⁴ die Einzelfallverhältnisse so realistisch wie möglich und ohne verzerrende Extremwerte abbildet.

⁵⁶⁸ Anwendungserlass zu § 158 AO: „Die Vorschrift enthält eine gesetzliche Vermutung. Sie verliert ihre Wirksamkeit mit der Folge der Schätzungsnotwendigkeit nach § 162 AO, wenn es nach Verprobung usw. unwahrscheinlich ist, dass das ausgewiesene Ergebnis mit den tatsächlichen Verhältnissen übereinstimmt.“

⁵⁶⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

⁵⁷⁰ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16.

⁵⁷¹ Wähnert, „Die Quantilsschätzung – Schätzungsmöglichkeiten mit summarischen Methoden“, BBK 20/2008, 1083.

⁵⁷² Schumann/Wähnert, „Die Quantilsschätzung – Schätzung anhand eines objektivierten Leistungsfähigkeitsmaßes“, Stbg 12/2012, 535.

⁵⁷³ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16.

⁵⁷⁴ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 166 AO, Rz. 44: „Naturgemäß besteht bei der Schätzung nach Wahrscheinlichkeitsgrundsätzen eine Bandbreite möglicher Wertansätze (sog. Schätzungsrahmen). Soweit sich die Schätzung innerhalb des Rahmens bewegt, den der durch die FinBeh. unter Mitwirkung des

- ! „Zudem ist die Anwendung der Quantilsermittlung je nach Länge des Zeitraums [...] als willkürlich anzusehen. Erfasst die Quantilsermittlung nur einen kurzen Zeitraum und liegen in diesem Zeitraum zufällig besonders hohe oder besonders niedrige Werte, verzerrt dieser Umstand in erheblicher Weise die Schätzung.“⁵⁷⁵
- ➔ Der Quantilsschätzungs-Box-Plot bewertet die Monatsgrößen des gesamten Prüfungszeitraums (s. den Aufzählungspkte. 6 „Übertragung der Normalverteilungsverhältnisse“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.). Handelt es sich um „einen kurzen Zeitraum [...] mit] zufällig besonders hohe[n] oder besonders niedrige[n] Werte[n]“, ist eine entsprechende Festlegung der Schätzungsgrößen sachgerecht. Sollten die „besonders hohe[n] oder besonders niedrige[n] Werte“ Ausreißer darstellen, werden sie über den Box-Plot aus dem Schätzungsrahmen ausgeschlossen. Durch die Einordnung der (betriebseigenen) Monatswerte mit relativen Prozenträgen (Quantilen) in Normalwerte und Ausreißer wird erreicht, dass das Ergebnis auch für unterschiedlich lange Prüfungszeiträume vergleichbar ist (s. den Aufzählungspkte. 3 „Rangbetrachtungen“ im Abschnitt „Annahmen und Arbeitsschritte“ unter diesem Pkt.).

Die Quantilsschätzung kann bei vorliegender Amtspflicht der Vollschatzung – **genauso wenig wie jede andere Methode** – die wahren Besteuerungsgrundlagen wiederherstellen. Sie bedeutet aber ein **verallgemeinerbares** Verfahren, um aus den Betriebsdaten einen Schätzungsrahmen abzuleiten, in welchem sehr wahrscheinlich die richtigen Bemessungsgrundlagen für die Steuern liegen. Die **bisherigen Erfahrungen** mit dem 80 %-Quantil aus mehreren tausend Schätzungsfällen seit 2008 haben ergeben, dass vernünftige, eher vorsichtige Schätzungsergebnisse resultieren. Entsprechend haben auch Überprüfungen der Quantilsschätzung durch Fahndungsmaßnahmen gezeigt, dass die tatsächlichen Steuerhinterziehungen in (nahezu) allen Einzelfällen vergleichbar oder sogar höher waren.

4.9. Strafrechtliche Schätzung

Das Verifikationsprinzip des § 158 AO (s. [Pkt. 4.6.](#)) mit seinem abgeschwächten Beweismaß „Beanstandungsanlass“ gilt nicht für das (Steuer-)Strafverfahren. Vielmehr muss die Finanzbehörde (Steuer-)Delikte für eine strafrechtliche Verurteilung beweisen. Dem Zweifelsgrundsatz „in dubio pro reo“ folgend setzt dies die volle richterliche Gewissheit voraus, weshalb „in

Stpfl. festgestellte Sachverhalt vorgibt, ist sie nicht zu beanstanden (BFH BStBl. 93, 259; 01, 381, 382; BFH/NV 99, 741 f.; 06, 1801, 1803; FG BW EFG 97, 45 f.; FG Hess. EFG 01, 798 f. m. Anm. Fumi). Der Schätzungsrahmen ist umso größer, je ungesicherter das Tatsachenmaterial ist, auf dem die Schätzung basiert.“

⁵⁷⁵ FG Berlin-Brandenburg v. 24.8.2016 – 5 V 5089/16.

solchen Fällen von der für den Angeklagten günstigsten Möglichkeit auszugehen [ist], die nach den gesamten Umständen in Betracht kommt“⁵⁷⁶.

Die Ausführungen des BGH zur ausreichenden Aussagegewissheit i.R. einer strafrechtlichen Bestimmung des Schuldumfangs „die ermittelte Vertrauensuntergrenze [...] mit einer Wahrscheinlichkeit von 99,5 % [als] zutreffend festgestellt“⁵⁷⁷ können als Anhaltspunkt für eine verallgemeinernde⁵⁷⁸ Einstufung des **Beweises** i.S. der vollen richterlichen Überzeugung herangezogen werden (s.a. [Pkt. 4.6.](#)). (Meta-)Studien zu Folge können auch moderne Beweismittel wie DNA- und Fingerabdruckanalysen keine höhere Gewissheit als 99-99,5 % gewährleisten.⁵⁷⁹ In Anbetracht der Tatsache, dass sich insbesondere komplexe(re) Vorgänge nicht vollständig und objektiv aus der Vergangenheit rekonstruieren lassen⁵⁸⁰, weil Dokumentationen, Aussagen etc. nur in seltenen Fällen vollständig und objektiv sind, erscheint diese Wahrscheinlichkeitsstufe zur Erlangung der vollen richterlichen Gewissheit ausreichend. Entsprechend konkretisieren die folgenden Hinweise des BGH eine realistische Einordnung des strafrechtlich erforderlichen Beweismaßes: „Liegen mehrere Beweisanzeichen vor, so genügt es nicht, sie jeweils einzeln abzuhandeln; erforderlich ist vielmehr eine Gesamtwürdigung. Auch wenn keine der jeweiligen Indiztatsachen für sich allein zum Nachweis der Täterschaft des Angekl. ausreicht, besteht die Möglichkeit, daß sie in ihrer Gesamtheit dem Gericht die entsprechende Überzeugung vermitteln können [...] Außerdem ist der Tatrichter nicht verpflichtet, entlastende Angaben des Angekl., für deren Richtigkeit oder Unrichtigkeit es keine Beweise gibt, ohne weiteres als unwiderlegt hinzunehmen. Vielmehr muß er sich auf der Grundlage des gesamten Beweisergebnisses entscheiden, ob diese Angaben geeignet sind, seine Überzeugungsbildung zu beeinflussen“⁵⁸¹.

Aus den mit Hilfe von Wahrscheinlichkeitstests abgesicherten Musteranalysen geht hervor, dass eine solche Aussagegewissheit von „≥ 99 bzw. 99,5 %“ zur Manipulation von Steuerungsdaten erreichbar ist (s. [Pkt. 4.3.1.3.](#)), wenn andere denkbare Ursachen für beo-

⁵⁷⁶ BGH v. 5.2.2003 – 2 StR 321/02.

⁵⁷⁷ Urt. v. 14.12.1989 – 4 StR 419/89, BGHSt 36, 320, Rz. 26.

⁵⁷⁸ Köller/Nissen/Rieß/Sadorf, „Probabilistische Schlussfolgerungen in Schriftgutachten – Zur Begründung und Vereinheitlichung von Wahrscheinlichkeitsaussagen in Sachverständigengutachten“, Reihe „Polizei + Forschung“, Bd. 26, herausgegeben vom Bundeskriminalamt (BKA), Kriminalistisches Institut, Luchterhand Verlag, 2004, 13: „Die numerische Schätzung bewirkt insbesondere eine Vermeidung von Missverständnissen oder möglichen Unterschieden in den scheinbar gleichen Größenordnungen verschiedener Individuen bei identischen verbalen Ausdrücken. [...] Sie erfordert vom Sachverständigen schlicht eine präzisere Einschätzung. Sie erhöht die intersubjektive Beurteilbarkeit dessen, was gemeint ist.“

⁵⁷⁹ Ausführlich hierzu: Giezek/Rupprecht/Wähnert, „Wie sicher muss ein Verprobungsergebnis sein? – Anleihen aus der Kriminalistik zur sinnvollen Einordnung des Beanstandungsanlasses“, BBK 2017, 236.

⁵⁸⁰ Vester, „Die Kunst vernetzt zu denken“, dtv Verlag, 2012, 15: „Unser Dilemma im Umgang mit der Komplexität unserer Welt lässt sich darauf zurückführen, dass wir wohl darin ausgebildet wurden einfache logische Schlüsse zu ziehen und nahe liegende Ursache-Wirkungs-Beziehungen zu definieren. [...] Simple Ursache-Wirkungs-Beziehungen gibt es nur in der Theorie, nicht in der Wirklichkeit.“

⁵⁸¹ BGH v. 25.11.1982 – 4 StR 564/82, NSTZ 1983, 133.

bachtete Auffälligkeiten (weitest gehend) ausgeschlossen werden können (s. [Pkt. 4.3.2.2.3.](#), [Pkt. 4.3.3.2.3.](#) u. [Pkt. 4.3.4.5.](#)). Diese Einschätzung entspricht der Beurteilung dieser Verfahren durch die Steuerstrafrechtskommentierung.⁵⁸² Zusätzlich sind bei Ergebnissen aus Prüfungsnetzen wie der SRP die Auswirkungen auf **grundsätzliche Beweisprinzipien** zu beachten:

- ➔ Infolge des systematisch-mehrperspektivischen Nachverfolgens von Auffälligkeiten in Besteuerungsdaten (s. [Pkt. 4.2.2.](#)) werden mögliche nicht-manipulative Ursachen (weitest gehend) abgeklärt und ggf. interaktiv berücksichtigt (s. [Pkt. 4.2.9.](#)). Dazu können alle konkreten Erkenntnisse aus der Mitwirkung und strafprozessualen Maßnahmen herangezogen werden. Verbleibende Unschlüssigkeiten im Gesamtbild erlangen somit eine (sehr) hohe Aussagegewissheit (s. den Abschnitt „*einzelne verbleibende Beweisanzzeichen*“ unter [Pkt. 4.4.](#)).
- ➔ Zusätzlich besteht ein beanstandetes Gesamtbild aus einem Prüfungsnetz wie der SRP erfahrungsgemäß aus mehreren (unabhängigen) Indizien, die sich als Indiziengesamtheiten⁵⁸³ in ihrer Aussageklarheit (erheblich) verstärken (s. den Abschnitt „*mehrere Indizien gegen die sachliche Richtigkeit*“ mit einem Beispiel unter [Pkt. 4.4.](#)).

Entsprechend wird in der Literatur zwischenzeitlich erkannt: „Der weitere Vorteil von SRP besteht darin, dass die finanzamtliche Beweisführung für die Hinzuschätzung auch im Steuerstrafverfahren gerichtsfest ist.“⁵⁸⁴

Auch bei der Feststellung des **Umfangs der (Steuer-)Straftaten** als Ausgangspunkt für das Strafmaß gilt der Zweifelsgrundsatz: „Im Strafverfahren genügt eine Wahrscheinlichkeit nicht, stattdessen wird die sichere Überzeugung über die Höhe der geschätzten Besteuerungsgrundlagen gefordert.“⁵⁸⁵ Andererseits geht „[d]er ‚In-dubio-pro-reo‘-Grundsatz [...] aber nicht soweit, dass eine mathematische Gewissheit für eine Verurteilung erforderlich wäre.“⁵⁸⁶

⁵⁸² Kohlmann, Steuerstrafrecht, 46. Lieferung 07.2012, § 397 AO, Rz 12.4: „Die praktische Verwertbarkeit der statistischen Verfahren ist durchaus umstritten, und zwar bereits als Grundlage für steuerliche Schätzungen, aber erst recht für strafrechtliche Verurteilungen. Die Finanzgerichte lehnen bislang den Chi-Quadrat-Test als eine sichere Methode, die Ordnungsgemäßheit der Buchführung zu verwerfen, ab. Allerdings wird weniger die grundsätzliche Eignung der Analysen angezweifelt als vielmehr hinterfragt, unter welchen Voraussetzungen und mit welcher Anzahl von Ausgangswerten die Methode aussagekräftige Ergebnisse liefert.“

⁵⁸³ BGH v. 27.4.2010 – 1 StR 454/09, NStZ 2011, 108, Rz. 18 u. 20: „Rechtsfehlerhaft ist es auch, wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern und auf ihren jeweiligen Beweiswert zu prüfen, ohne eine Gesamtabwägung aller für und gegen die Täterschaft sprechenden Umstände vorzunehmen [...] Denn die Indizien können in ihrer Gesamtheit [...] die entsprechende Überzeugung vermitteln, auch wenn [sie] jeweils für sich allein nicht zum Nachweis der Täterschaft des Angeklagten ausreich[en].“

⁵⁸⁴ Pump/Heidl, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenbuchführung“, StBp 2014, 163.

⁵⁸⁵ Brinkmann, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017, 481.

⁵⁸⁶ Brinkmann, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017, 481 f.

Deshalb sind Schätzungen der hinterzogenen Steuern auch im Strafverfahren zulässig⁵⁸⁷ und „notwendig, wenn keine oder nur mangelhafte Aufzeichnungen vorhanden sind. Denn ein Freispruch kommt nicht in Betracht, wenn die Schuld des Täters als solche feststeht“⁵⁸⁸. Demzufolge sieht *Brinkmann* eine „[verringerte] Zurückhaltung der strafgerichtlichen Rechtsprechung gegenüber dem Thema ‚Schätzung‘ [...] in den letzten Jahren“⁵⁸⁹.

Die **Quantilsschätzung** (s. [Pkt. 4.8.](#)) verfolgt streng die Maßgabe einer Schätzung im Besteuerungsverfahren, dass „der ‚Beweisverderber‘ oder ‚Beweisvereitler‘ aus seinem Verhalten keinen Vorteil ziehen darf (Rechtsgedanke des § 444 ZPO i.V.m. § 155 FGO)“⁵⁹⁰. Danach „ist das FA im allgemeinen nicht nur berechtigt, sondern verpflichtet, in dem gegebenen Schätzungsrahmen an die oberste Grenze zu gehen.“⁵⁹¹ Mit dieser Ausrichtung ist sie nicht (direkt) für eine strafrechtliche Schätzung geeignet.

Demgegenüber rekonstruiert die **Verteilungsanalyse** mit ihrer quantitativen Auswertung (s. [Pkt. 4.3.4.6.](#)) die wahrscheinliche Nachfragesituation des geprüften Stpfl. aus dessen (betriebseigenen) Tageskasseneinnahmen und in Anlehnung an wissenschaftliche Verteilungserkenntnisse (s. [Pkt. 4.3.4.1.](#)). Sofern das typische Manipulationsbild „Spitzeneingriff“ (s. [Pkt. 4.3.4.3.](#)) erkennbar ist und nicht-manipulative Ursachen (weitest gehend) ausgeschlossen werden können (s. [Pkt. 4.3.4.5.](#)), beziffert das Ergebnis den Mindestschaden. „Wenn der BGH – unter den von ihm vorgegebenen Voraussetzungen – im Rahmen der tatrichterlichen Beweiswürdigung selbst so generalisierende Schätzungsmethoden wie die Schätzung nach der Richtsatzsammlung zulässt, dann sollten auch die Ergebnisse der statistischen-mathematischen Methoden für die richterliche Überzeugungsbildung von Relevanz sein, da der Beweiswert eines inneren Betriebsvergleichs [...] sicherlich höher einzuschätzen ist als ein äußerer Betriebsvergleich (Richtsatzsammlung).“⁵⁹² Deshalb könnte diese Methode auch als strafrechtliches Schätzungsverfahren in Betracht kommen, wenn keine konkreteren Informationen vorliegen.

⁵⁸⁷ BGH v. 4.2.1992 – 5 StR 655/91, wistra 1992, 147 u. v. 24.9.1985 – 1 StR 313/85, wistra 1986, 65.

⁵⁸⁸ *Brinkmann*, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017, 482.

⁵⁸⁹ *Brinkmann*, „Strafrechtliche Aspekte der Schätzung“, StBp 9/2013, 250, 253.

⁵⁹⁰ BFH v. 15.2.1989, BStBl. II 1989, 462: „[...] Zur Vermeidung eines solchen Ergebnisses sind auch belastende Unterstellungen [...] oder nachteilige Schlüsse im Rahmen der Beweiswürdigung gerechtfertigt“.

⁵⁹¹ BFH v. 9.3.1967, BStBl. III 1967, 349.

⁵⁹² *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24, 25.

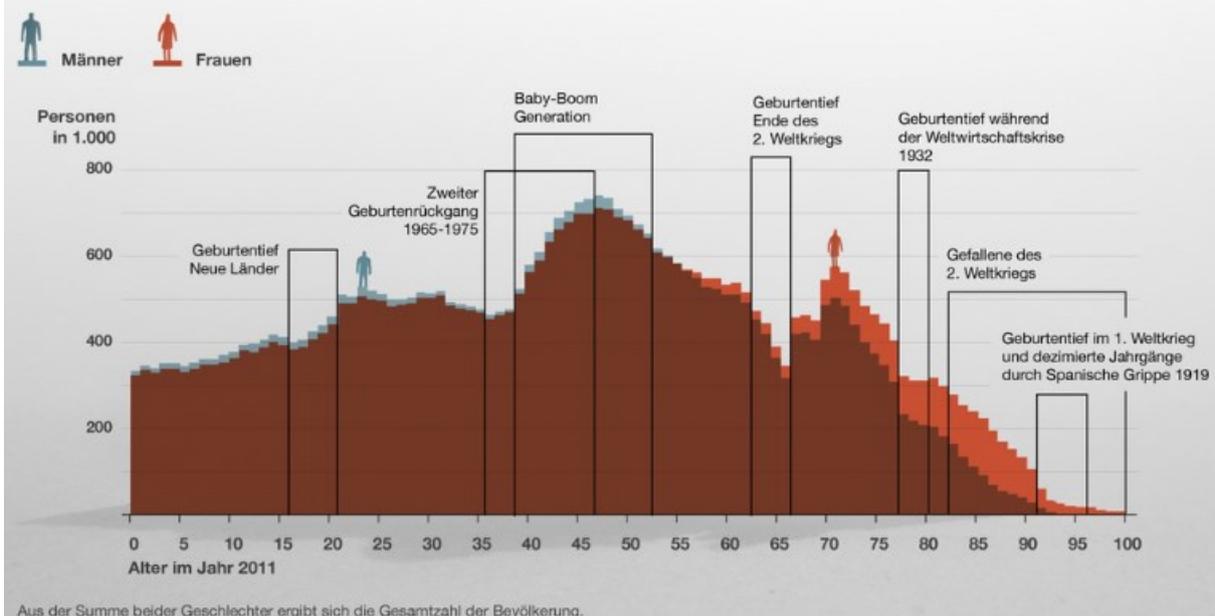
5. Freie SRP-Anwendung (Systematische Visualisierung)

Die **SRP-Vorlagen sind Einstiegshilfen** in eine systematisch-mehrperspektivische Datenvisualisierung für ein effektiveres Erkennen von betriebsspezifischen Prüffeldern und ein effizienteres Abarbeiten mit aussagekräftigeren Ergebnissen. Sie sollen besonders dem Teil des Prüfungspersonals mit wenig IT-Affinität einen Einblick in die Vorteile der verbildlichenden Prüfung geben, ohne dass dafür umfangreiche IT-Kenntnisse im Vorfeld angeeignet werden müssen.

Wie viel leistungsfähiger (systematisch-mehrperspektivische) Datenvisualisierungen gegenüber den herkömmlichen Herangehensweisen in der Prüfung sind, lässt sich nicht nur aus den Erfahrungen der Werbebranche⁵⁹³ und aus Wissenschaftsbereichen mit besonders komplexen Fragestellungen ableiten wie bei dieser Verbildlichung der demografischen Verhältnisse in Deutschland⁵⁹⁴:

Die Altersstruktur Deutschlands nach Geschlecht 2011

Die jüngsten Jahrgänge sind zahlenmäßig kleiner als viele der älteren, obwohl die Zahl der Angehörigen eines Jahrgangs mit zunehmendem Alter abnimmt. Ursachen sind die gesunkene Geburtenrate und auch die verringerte Sterblichkeit der Älteren. Zudem veranschaulicht die Anzahl der Frauen und Männer in den jeweiligen Jahrgängen die Auswirkung historischer Ereignisse auf die Zusammensetzung der Bevölkerung.



⁵⁹³ *Barnard* in einer US-amerikanischen Zeitschrift der Werbebranche 1921: „One look is worth a thousand words“, nachzulesen unter: www.oai.de/de/component/content/article/45-publikationen/sprichwort/811-ein-bild-sagt-mehr-als-tausend-worte.html.

⁵⁹⁴ Bildquelle: Internetseite der Bundeszentrale für politische Bildung (bpb), Bereich „Politik/Innenpolitik/Demografischer Wandel“, nachzulesen unter: <http://www.bpb.de/politik/innenpolitik/demografischer-wandel/209140/infografiken>; Lizenz: [cc by-nc-nd/3.0/de/](http://creativecommons.org/licenses/by-nc-nd/3.0/de/).

Auch im Eigenversuch an einem einfachen Beispiel *im Abschnitt „Visualisierung von Steuerungsdaten“ unter [Pkt. 2.](#)* lassen sich deren Vorteile nachvollziehen. Tlw. ist diese Kenntnis in den prüfenden Berufen durchaus seit langer Zeit bekannt, wenn Unternehmensverbände in Form von Organigrammen anstatt als Tabelle wiedergegeben werden. Umso erstaunlicher ist es, dass die viel komplexeren Steuerungsdaten noch verbreitet in Tabellenform geprüft werden. Inzwischen werben aber immer mehr professionelle Hersteller von Datenanalyse-Software mit ihren Visualisierungsmöglichkeiten.⁵⁹⁵

Da die Fragestellungen der Prüfungspraxis – abgesehen von einigen Prüffeldern, die weitgehend ähnlich gelagert sind wie die Vollständigkeit von (Bar-)Einnahmen, – oft einzelfall-spezifisch sind, kann eine hohe (digitale) Prüfungskompetenz nur durch die Fähigkeiten zur **freien Visualisierung** i.S. folgender Erkenntnis aus der Wissenschaft erlangt werden: „Die erfolgreiche Durchführung betriebswirtschaftlicher Ordnungsmäßigkeitsprüfungen erfordert von den Beteiligten ein hohes Maß an Methodenwissen, Flexibilität und Kreativität.“⁵⁹⁶ Mit dem SRP-Konzept – im Gegensatz zu den vollautomatischen Analysesystemen der künstlichen Intelligenz⁵⁹⁷ – wird dazu das grundlegende Methodenwissen in den Vordergrund gestellt (s. den Abschnitt „Methodenwissen“ mit einem Beispiel unter [Pkt. 2.](#)). Mit der zielgerichteten Verknüpfung diverser *geeigneter* Verbildlichungsmethoden unterstützen die SRP-Vorlagen eine variable und schnelle Arbeitsweise in vielen Bereichen wie bspw. auch der Schnittstellenprüfung (s. [Pkt. 5.3.](#)), so dass die einzelfallabhängige Kombination aus vorlagengestützter und freier SRP-Anwendung i.V. mit anderen Ansätzen wie z.B. mathematisch-statistischen Stichproben (s. den Abschnitt „statistische Stichprobenverfahren“ unter [Pkt. 2.](#)) die größte Verbesserung bewirkt.

5.1. Einsatzbereiche systematischer Visualisierung

Weil gerade bei der anfänglich fehlenden Erfahrung mit neuartigen Prüfungsansätzen noch viel Unsicherheit darüber besteht, in welchen Bereichen sie erfolgversprechend eingesetzt werden können, soll die folgende **nicht abschließende Übersicht** dabei helfen:

⁵⁹⁵ Vgl. z.B.: Internetblog der Fa. Audicon zu IDEA V10: „Grafische Darstellungen lassen sich im Allgemeinen leichter erfassen als die Darstellung der Daten in Tabellenform. Daher ist ein besonderes Feature des neuen Releases die Visualisierung“, nachzulesen unter: <https://blog.audicon.net/audit-analytics/ideav10-visualisierung-und-ermitteln.html> oder Visualisierungsprogramm „Visual Analytics“ des amerikanischen, weltweit operierendes Softwarehauses Statistical Analysis Systems, nachzulesen unter: http://www.sas.com/de_de/software/business-intelligence/visual-analytics.html.

⁵⁹⁶ Nachzulesen unter: www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik.

⁵⁹⁷ *Kompenhans/Wermelt*, „Big Data“ – Potenziale für innovative Abschlussprüfungen“, BB 2018, 299: „Die an Systemschnittstellen erforderlichen manuellen Tätigkeiten werden immer mehr durch den Einsatz von Robotic Process Automation ersetzt, Entscheidungsprozesse werden dabei durch den Einsatz programmierter, selbstlernender Entscheidungsheuristiken, die auf Basis künstlicher Intelligenz arbeiten, unterstützt.“



Nur durch wiederholtes Anwenden (mindestens) in diesen Bereichen kann die notwendige Erfahrung und Übung mit systematisch(-mehrperspektivisch)er Datenvisualisierung entstehen, um die ganzen Vorteile in der Prüfung genießen zu können. Mit der Zeit resultieren die „[erforderliche] Flexibilität und Kreativität“⁵⁹⁸ (s. [Pkt. 5.](#)), um aus komplexen oder schwierigen Prüfungsfragen auf Basis der verfügbaren (digitalen) Daten und mit Hilfe der zur Verfügung stehenden technischen Mittel eigene Verbildlichungen zielführend und verständlich zu erstellen. Dabei ist der (zeitliche) Aufwand i.d.R. verhältnismäßig gering.

Heutzutage hat sich „[d]ie Buchführung [...] von der Durchschreibebuchführung hin zu einer umfassenden Datenverarbeitung entwickelt, die elektronische Form ist nunmehr die übliche. [...] Dabei steigt die Anzahl elektronisch überwachter Prozesse stetig.“⁵⁹⁹ In der Folge ist es selbst bei kleinen Betrieben mit (tlw.) digitaler Aufzeichnung nicht mehr möglich, eine (weitgehende) Vollprüfung zu erreichen (s. [Pkt. 2.1.](#)). Stichprobenprüfungen können „die Ermittlung und Beurteilung der steuerlich bedeutsamen Sachverhalte“⁶⁰⁰ nicht gewährleisten. Deshalb ist es inzwischen unerlässlich, dass für eine zielführende, ressourcenschonende Durchführung zu Beginn einer Außenprüfung ein **systematischer Überblick** über die Besteuerungsdaten erarbeitet wird. Hierbei ist die Datenvisualisierung aufgrund ihrer wahrnehmungsphysiologischen Vorteile (s. [Abschnitt „Visualisierung von Besteuerungsdaten unter Pkt. 2.](#)) ein wichtiger Bestandteil (s. [Pkt. 5.2.](#)).

Wesentliches Argument für die zunehmende Automatisierung der Betriebsverwaltungen⁶⁰¹ war und ist die Einsparung von Personal. Gemäß der GoBD „[hat der Steuerpflichtige

⁵⁹⁸ Nachzulesen unter: www.wpunco.wiwi.uni-due.de/forschung/pruefungstechnik.

⁵⁹⁹ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876 u. 877.

⁶⁰⁰ § 2 BpO: „Aufgaben der Betriebsprüfungsstellen: Zweck der Außenprüfung ist die Ermittlung und Beurteilung der steuerlich bedeutsamen Sachverhalte, um die Gleichmäßigkeit der Besteuerung sicherzustellen.“

⁶⁰¹ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 876: „Die Buchführung hat sich von der Durchschreibebuchführung hin zu einer umfassenden Datenverarbeitung entwickelt, die elektronische Form ist nunmehr die übliche. [...] Diese

f]ür die Einhaltung der Ordnungsvorschriften des § 146 AO [...] Kontrollen einzurichten, auszuüben und zu protokollieren.“⁶⁰² Dass dies (noch) nicht ausreichend geschieht, bestätigen sogar die Spezialisten in den Unternehmen.⁶⁰³ Aus regelmäßigen Prüfungserfahrungen ist bekannt, dass Unstimmigkeiten in den Betriebsdaten nicht nur infolge von Migrationen auftreten sondern auch durch die Einstellung und Bedienung der komplexen Systeme und offenbar sogar innerhalb der automatisierten Prozesse. Darum wird schon länger in der Literatur darauf hingewiesen, dass die **Untersuchung von Datenschnittstellen** ein heutiges Schwerpunktprüffeld sein sollte.⁶⁰⁴ Auch hier bietet die visualisierte Gegenüberstellung der Daten an der Schnittstelle eine schnelle und wirksame Lösung (s. [Pkt. 5.3.](#)).

Neben der Prüfung von Einzelsachverhalten hat die **Schlüssigkeitsprüfung** von Gesamtbildern in der Besteuerungsverifikation eine wichtige Bedeutung.⁶⁰⁵ Bei datenintensiven oder evtl. verschleierte Sachverhalten, für welche die Finanzverwaltung allgemein die Beweislast trüge, gilt dabei nach § 158 AO das besondere Beweismaß „Beanstandungsanlass“ (s. [Pkt. 4.6.](#)). Hiervon ist nicht nur die Überprüfung erklärter (Bar-)Einnahmen auf Vollständigkeit betroffen sondern bspw. auch der Prüfungsgegenstand, ob Verträge zwischen nahestehenden Personen steuerlich anzuerkennen sind. Insbesondere bei der (betriebswirtschaftlichen) Betrachtung der Fremdüblichkeit von konzerninternen Verrechnungen hilft die Verbildlichung der tatsächlichen Verhältnisse auf Basis der dazugehörigen Einzeldaten häufig erheblich, um zunächst eine sachgerechte Einschätzung zu erlangen und anschließend um diese auf einem transparenten Ermittlungsergebnis zu besprechen (s. [Pkt. 5.4.](#)).

Entwicklung geschah auf Betreiben und im Interesse der Unternehmen, weil es betriebswirtschaftlich notwendig war und ist, diese Einzeldaten zu erzeugen, z.B. für ein aussagekräftiges Controlling oder zur Arbeitsprozesssteuerung.“

⁶⁰² „Grundsätze zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)“, BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450, Rz. 100: „6. Internes Kontrollsystem (IKS)“.

⁶⁰³ Vgl. z.B.: Newsletter 4-2016 des Forums „Elektronische Steuerprüfung“ v. 22.4.2016: „(IT-)Management: Ein Drittel aller Unternehmen verlieren Daten bei der Migration: Ein Drittel (32 %) der Organisationen verlieren Daten, während sie Daten zwischen Geräten migrieren oder Betriebssysteme aktualisieren. Das ist das Ergebnis einer weltweiten Befragung unter fast 600 IT-Administratoren“, nachzulesen unter: http://elektronische-steuerpruefung.de/newsarchiv/news_16_04.htm?m=1450705466#Beitrag7.

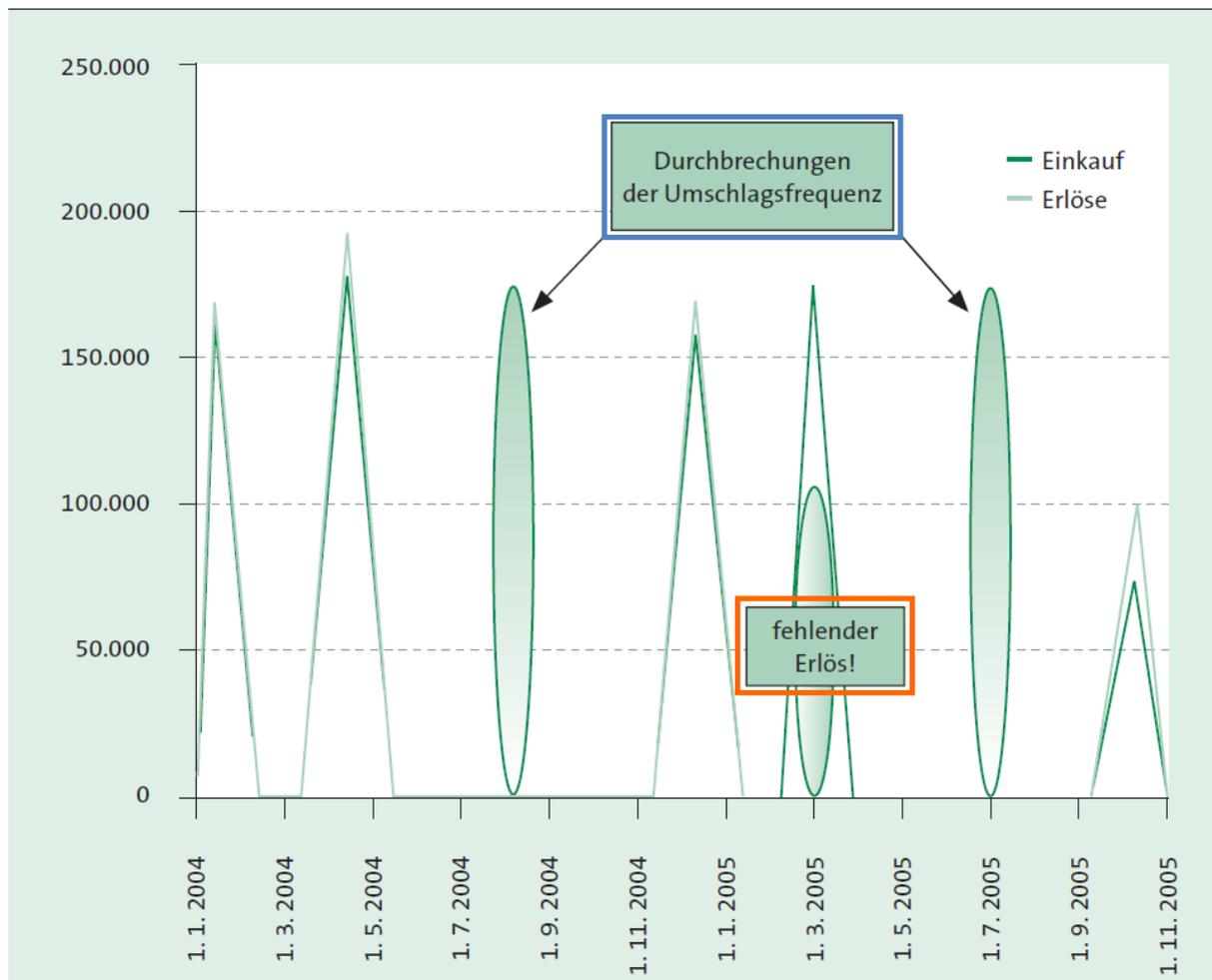
⁶⁰⁴ Vgl. z.B.: *Webel/Danielmeyer*, „Schnittstellenverprobung elektronischer Betriebsverwaltungen – das kommende Standardprüffeld“, StBp 2015, 353, sowie *Danielmeyer/Neubert/Unger*, „Praxiserfahrungen zu Vorsystemen – Prüffeld: Visualisierte Schnittstellenverprobung“, StBp 2016, 322.

⁶⁰⁵ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 5: „Die gesetzliche Regelung des § 158 gründet sich insofern auf die Einsicht, dass es weder möglich noch wünschenswert ist, eine Unzahl von Geschäftsvorfällen lückenlos nachzuprüfen, und deshalb das äußere Gesamtbild einer Buchführung als ausreichende Grundlage für die Sachverhaltsfeststellung genügen kann.“

5.2. Visualisierende Prüffeldsuche

Die chronologische Nachverfolgung von Periodensummen aus Unternehmensdaten zur Beobachtung von Entwicklungen und Abhängigkeiten entspricht der Idee der betriebswirtschaftlichen Analysen (BWA), beachtet dabei aber die wahrnehmungsphysiologischen Vorteile der Verbildlichung (s. den Abschnitt „Visualisierung von Besteuerungsdaten“ unter [Pkt. 2.](#)). Wird bei derartigen **regelbasierten Ansätzen** zudem gem. den Fuzzy-Theorien (s. den Abschnitt „Fuzzy-Theorien“ unter [Pkt. 2.](#)) berücksichtigt, wie sehr der Grad der Aufsummierung von Daten (Summationsebene) für das mögliche Erkennen von unterschiedlichen Auffälligkeiten von Bedeutung ist, kann mit wenig Aufwand ein ergebnisoffener und tiefreichender Einblick in die Verhältnisse eines zu prüfenden Betriebs erarbeitet werden.

In der beispielhaft abgebildeten Zeitreihenanalyse wurde dieser Einstieg in die Prüfung eines Großhandelsunternehmens gewählt:⁶⁰⁶



Für einen der Warenbereiche, die vorab über die Kostenstellen aufgetrennt worden waren, ergaben sich die im Diagramm markierten Auffälligkeiten. Der scheinbar **fehlende Erlös** wur-

⁶⁰⁶ Bildquelle: Weiskopf/Wähner, „Neue interaktive Prüfungstechnik – Transfereffekte moderner Verprobungsansätze auf die Großbetriebsprüfung“, BBK 3/2012, 127, 129.

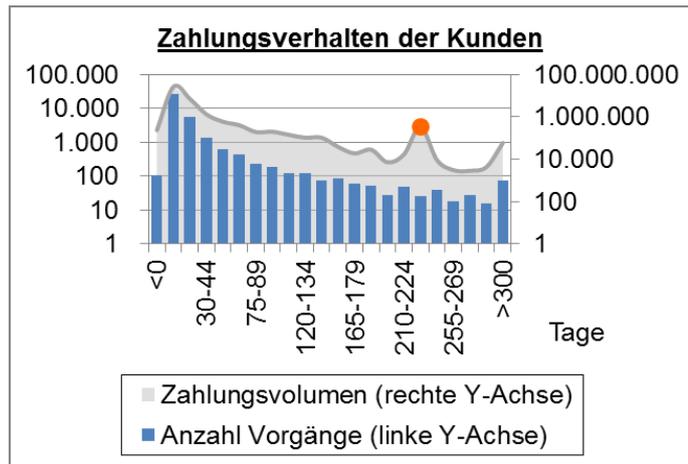
de zügig als Erfassungsunregelmäßigkeit auf einem anderen Konto wiedergefunden. Demgegenüber führte die vertiefende Auseinandersetzung mit der offensichtlichen **Durchbrechung der typischen Umschlagsfrequenz** zu einem unerwarteten Prüffeld: „Darüber hinaus sind die oszillierenden Einkaufs- und Verkaufsamplituden von Interesse, die das Produkt der betrieblichen Rahmenverhältnisse sind, zu denen sowohl die Lieferkapazitäten der Erzeuger als auch das Nachfrageverhalten der Kunden zählen. Bei dieser Produktgruppe ist die Tatsache bedeutsam, dass sie neben dem Wareneinsatz mit ergänzendem Provisionsaufwand belastet ist. Weitere durch das Verprobungsergebnis veranlasste Untersuchungen lieferten die Erkenntnis, dass es in der Belieferung künstliche Engpässe gab, um den Provisionsforderungen Nachdruck zu verleihen. Somit kann diese Art der Analyse auch für das Fraud-Management hilfreich sein.“⁶⁰⁷ Als steuerliches Ergebnis stellte sich heraus, dass im Zusammenhang mit den Provisionsstreitigkeiten erhebliche nicht abzugsfähige Betriebsausgaben gem. § 4 Abs. 5 Nr. 10 EStG gewinnwirksam erfasst wurden.

Werden **geeignete** (s. den Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) regelbasierte Untersuchungen an unterschiedlichen Unternehmensdaten vom Waren- bzw. Leistungsfluss über Kostenarten und -stellen bis zu Vertragsdurchführungen systematisch vom groben Überblick zu Detailperspektiven – entsprechend der **SRP-Zeitreihensystematik** – durchgeführt (s. den Abschnitt „Eingrenzung aufgefundener Regelstörungen zu ursächlichen Sachverhalten“ unter [Pkt. 4.2.2.](#)), können zielführend generelle und punktuelle Prüffelder des Einzelfalls identifiziert werden.

Musterbasierte Häufigkeitsauswertungen setzen auf der in der Wirtschaftsprüfung üblichen ABC-Analyse auf und beachten dabei ebenfalls die wahrnehmungsphysiologischen Vorteile der Verbildlichung (s. den Abschnitt „Visualisierung von Besteuerungsdaten“ unter [Pkt. 2.](#)). Mit Verteilungsmustern zu den ausgewerteten Merkmalen umfangreicher Unternehmensdaten (s. den Abschnitt „Auszählung von Informationen nach unterteilbaren Eigenschaften“ unter [Pkt. 4.3.](#)) arbeiten sie eine andere Information als regelbasierte Ansätze heraus. Insofern konkurrieren regel- und musterbasierte Analysen bei der Gewinnung eines Überblicks über umfangreiche Besteuerungsdaten nicht, sondern ergänzen sich. Im Vergleich der beobachteten Häufigkeitsmuster mit betriebsinternen Vergleichsdaten oder allgemeinen (externen) Referenzen wie der Logarithmische Normalverteilung (s. [Pkt. 4.3.1.2.](#) u. [Pkt. 4.3.4.1.](#)) ist ebenso auf die Fuzzy-Theorien (s. den Abschnitt „Fuzzy-Theorien“ unter [Pkt. 2.](#)) zu achten, indem die Intervallgröße bei der Klassenbildung (s. den Abschnitt „Klassenunterteilung mit Wurzelmethode“ unter [Pkt. 4.3.3.1.](#)) darüber mitentscheidet, welche Auffälligkeiten erkennbar sind.

⁶⁰⁷ Weiskopf/Wähnert, „Neue interaktive Prüfungstechnik – Transfereffekte moderner Verprobungsansätze auf die Großbetriebsprüfung“, BBK 3/2012, 127, 129.

In dieser beispielhaften Häufigkeitsauswertung wurden die Debitorendaten eines Konzernunternehmens auf das Zahlungsverhalten aller Kunden hin untersucht, weil mögliche Vorzugsbehandlungen verbundener Unternehmen steuerlich evtl. als verdeckte Gewinnausschüttungen zu behandeln sind.⁶⁰⁸ Dazu wurden alle bezahlten Rechnungen in 15-Tages-Intervallen ausgezählt und zusammen mit dem Forderungsvolumen pro Klasse im Diagramm abgebildet. Für die beiden Y-Achsen wurde jeweils die logarithmische Form gewählt, um den Fokus auf einzelne Ausreißer als betriebsspezifische Prüffelder zu lenken. Die gleichmäßige Häufigkeitsabnahme unter zunehmender Zahlungsdauer ist bis zur vorletzten Klasse weitgehend gegeben, denn die Unruhen rechtsseitig von „165 bis 179 Tage“ bewegen sich im empfindlichen⁶⁰⁹ unteren Skalenbereich. Mit dem letzten Balken werden alle Zahlungsvorgänge ab einer Zahlungsferne von 300 Tagen zusammengefasst. Deshalb verwundert der leichte Anstieg nicht. Aus der zusätzlichen Information der klassenweise aufsummierten Forderungsvolumina sticht demgegenüber das Intervall „225 bis 239 Tage“ hervor: Trotz einer (passend) geringen Anzahl an Vorgängen ergibt sich ein vergleichsweise hoher Forderungsbetrag von mehr als € 300.000. Demnach müssen eine oder wenige Rechnungen entsprechend lange unbezahlt geblieben sein. Die gezielte Filterung der Daten aus der betreffenden Klasse und die anschließende Begutachtung der dazugehörigen Rechnungen ergab, dass es sich um wenige Rechnungen an ein verbundenes Unternehmen handelte und dass die Langzeitforderung nicht verzinst worden war. Neben dem Einzelaufgriff müssen keine weiteren (stichprobenartigen) Prüfungen der bezahlten Forderungen für das Unternehmen, aus dem die untersuchten Daten stammen, vorgenommen werden, weil alle anderen konzerninternen Verrechnungen im fremdüblichen Zahlungsverhalten unauffällig „untergehen“.

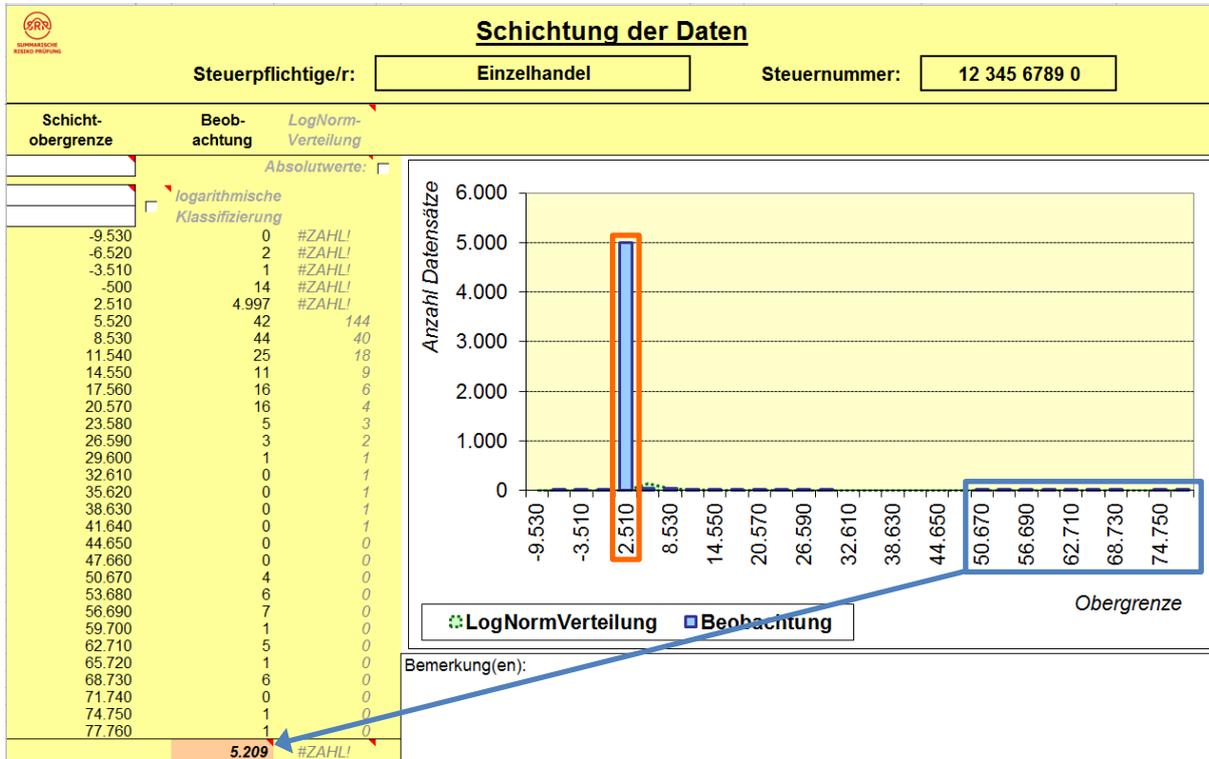


Die **SRP-Vorlage Stochastik** bietet mit den beiden Tabellenblättern „Schichtung“ und „Schichtungsvertiefung“ eine praxisgerechte Unterstützung, um große Mengen an Besteuerungsdaten *systematisch* auf Häufigkeitsmuster und prüfungsrelevante Abweichungen hin zu untersuchen. Der SRP-Prüfungssystematik „vom Groben ins Detail“ folgend ist in die-

⁶⁰⁸ Vgl. Weiskopf/Wähnert, „Neue interaktive Prüfungstechnik – Transfereffekte moderner Verprobungsansätze auf die Großbetriebsprüfung“, BBK 3/2012, 127, 131.

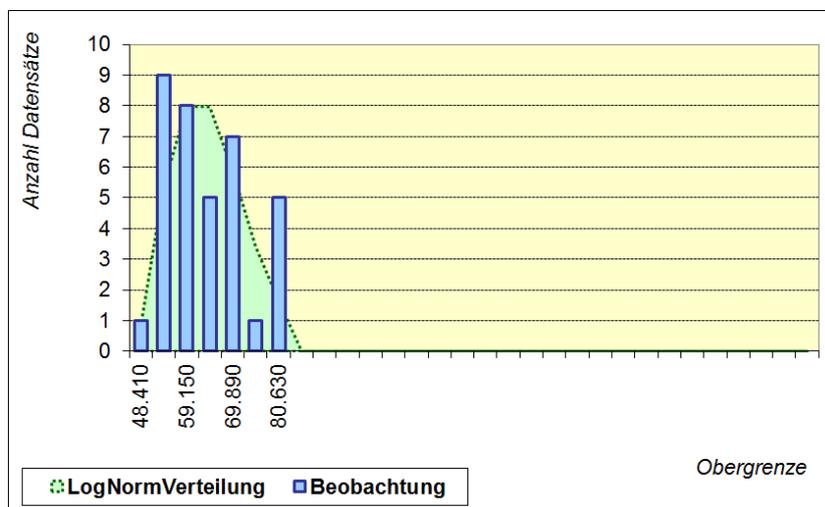
⁶⁰⁹ Die gleichmäßig beschrifteten Abstände der beiden logarithmischen (zu Basis „10“) Y-Achsen nehmen in ihrer Bedeutung von oben nach unten pro Intervall um den Faktor „10“ ab. Somit werden bspw. die Schwankungen um „210 bis 224 Tage“ und „255 bis 269 Tage“ herum 10-mal intensiver dargestellt als die Unruhen der Balken im darüber liegenden Häufigkeitsintervall „100 bis 1.000“.

ser ersten Übersichtsgrafik auf dem Tabellenblatt „Schichtung“ der wertabhängige Verteilungszusammenhang für die gesamten Kostenpositionen eines Einzelhandelsunternehmens zu sehen:



Wegen einer Gruppe von Höchstbeträgen, welche gem. Warnhinweis nicht vollständig angezeigt werden (s. den Auflistungspkt. „nicht alle hohen Spitzenwerte in den Auswertungen berücksichtigt“ unter [Pkt. 4.3.3.1.](#)), fallen nahezu alle Sachverhalte in eine Klasse. Außerdem können der nebenstehenden Tabelle drei Kostenpositionen mit erheblichen Negativbeträgen abgelesen werden, welche als Einzelprüffelder gefiltert und rechtlich geprüft werden können.

In einem ersten Vertiefungsschritt werden die außergewöhnlich hohen Aufwendungen

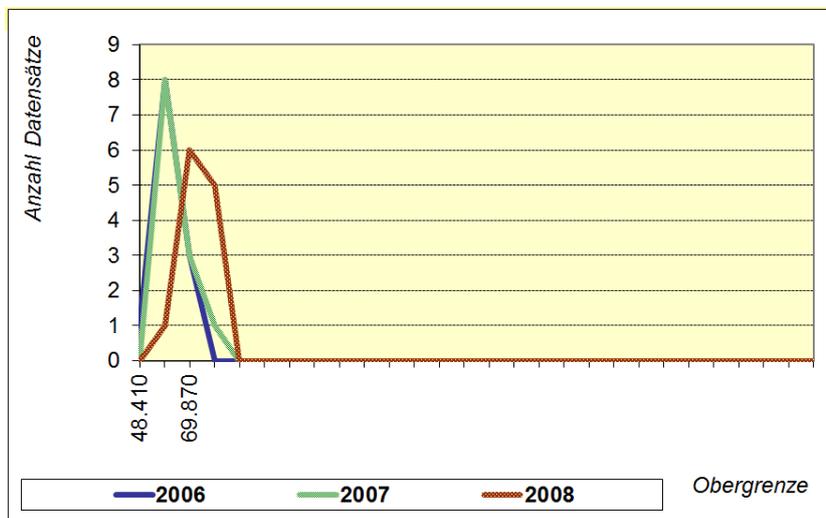


durch Eingabe eines entsprechenden Startwerts (s. den Auflistungspkt. „Startwert“ unter [Pkt. 4.3.3.1.](#)) isoliert betrachtet. Die insgesamt 36 Werte der Abb. stellen scheinbar einen eigenständigen Bestand an Monatssummen dar. Durch eine gezielte Datenfilterung der Buchführung

lässt sich umgehend herausfinden, dass es die Monatsabrechnungen eines Großhändlers

sind. Dabei ergibt sich aber kein gleichmäßiges Häufigkeitsmuster. Die Ursache hierfür soll mittels interner Gegenüberstellung der drei Prüfungsjahre auf dem Tabellenblatt „Schichtungsverteifung“ untersucht werden.

Dabei offenbart sich, dass die ersten beiden Jahre eine fast identische Häufigkeitsverteilung aufweisen, während die zwölf Monatsrechnungen des letzten Jahres insgesamt ein deutlich höheres Niveau erreichen. Sofern sich diese Veränderung nicht durch einen höheren



Gesamtabsatz oder den Umstieg von einem anderen Lieferanten nachvollziehbar begründen lässt, sollte z.B. stichprobenartig nachgeprüft werden, ob „Gefälligkeiten“ der Auslöser sind.

Weil auch die anderen Kostenpositionen im fünfstelligen Bereich aus monatlich zusammengefassten Wareneinkäufen von anderen Großhändlern stammen, werden die übrigen Aufwendungen durch eine schrittweise verfeinerte Schichtung (s. den Auflistungspkt. „freie Klassenbildung“ unter [Pkt. 4.3.3.1.](#)) auf deren Häufigkeitsmuster und prüfungsrelevante Datenkonzentrationen untersucht:

Schichtung der Daten

Steuerpflichtige/r: Einzelhandel Steuernummer: 12 345 6789 0

Untersuchte Daten: Einzeldatenabwahl

Schichtobergrenze	Beobachtung	LogNorm-Verteilung
1	1	0
101	2.285	2.234
201	723	721
301	395	393
401	337	254
501	222	179
601	91	134
701	101	104
801	68	84
901	59	69
1.001	58	57
1.101	39	49
1.201	42	42
1.301	30	36
1.401	22	32
1.501	20	28
1.601	5	25
1.701	17	22
1.801	12	20
1.901	4	18
2.001	39	16
2.101	21	15
2.201	2	14
2.301	13	12
2.401	8	11
2.501	8	11
2.601	4	10
2.701	2	9
2.801	6	8
2.901	2	8
4.636	4.615	

Bemerkung(en):

Durch den **Startwert „1“** werden nur die positiven Datensätze analysiert, wodurch die Logarithmische Normalverteilung als Orientierung vollständig berechnet werden kann⁶¹⁰. Eine

untersuchte Werte	lg	absolute Anzahl:
161,87	2,209	
182,62	2,262	

Benutzerdefinierter AutoFilter

Zeilen anzeigen:
 untersuchte Werte

ist größer als

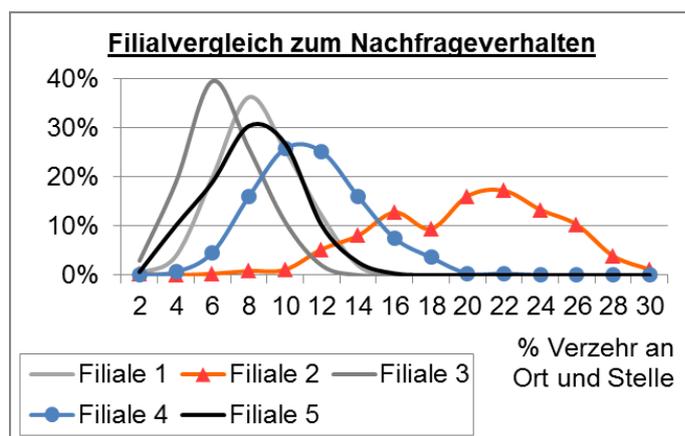
Und Oder

ist kleiner oder gleich

erste **Vorgangshäufung** ist in dem Intervall „> € 1.901 bis € 2001“ zu erkennen. Ob es sich dabei um eine vertragliche Festlegung von Periodenwerten oder ein Vertiefungsfeld handelt, kann über die **Filterung der betreffenden Datensätze** auf dem Tabellenblatt „Vergleichsdaten“ und die anschließende Untersuchung der Belege zielführend ermittelt werden. Über das Verkleinern des Klassifizierungsintervalls mit Hilfe der Eingabefelder für die ersten beiden Obergrenzen kann die Prüffeldsuche von starken zu schwächeren Datenkonzentrationen ausgeweitet werden.

Weil die chronologischen und die Verteilungsinformationen aus Besteuerungsdaten unterschiedliche Visualisierungsperspektiven bedeuten, die sich ergänzen, lässt sich durch die **kombinierte regel- und musterbasierte Datenverbildlichung** die effizienteste Prüffeldaufdeckung und -eingrenzung erreichen wie bei der Vertiefung einer beobachteten Zeitreihenauffälligkeit mit Hilfe einer musterbasierten Häufigkeitsauswertung *im Abschnitt „Prüfungsnetz“ unter Pkt. 2.3.*

Als Beispiel zur Bearbeitung eines konkreten Prüffelds soll die Trennung der Entgelte in einem Großbetrieb mit vielen Filialen überprüft werden. Um über die vielen Millionen Einzeldaten aus den einheitlichen Kassensystemen einen Überblick für die gezielte Vertiefung zu gewinnen, werden die prozentualen Tagesanteile „Verzehr an Ort und Stelle“ als spezifische Häufigkeitsmuster der einzelnen Filialen gegenübergestellt. Infolge der ja-nein-Entscheidung,



ob im Geschäft oder außerhalb verzehrt wird, müsste sich pro Filiale eine Binomialverteilung um das typische Nachfrageverhalten herum ausbilden. Zur besseren Übersichtlichkeit werden im nebenstehenden Diagramm nur fünf Filialen ausgewertet und die Häufigkeitsmuster als prozentuale Linien grafen⁶¹¹ wiedergegeben. Die Häufig-

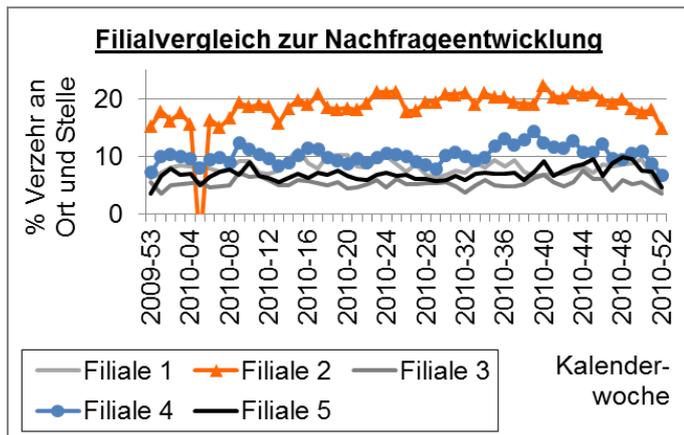
⁶¹⁰ Da es keinen Logarithmus zu Null und negativen Zahlen gibt, kann die Herleitung der betriebsspezifischen Logarithmischen Normalverteilung nur für den positiven Wertebereich erfolgen.

⁶¹¹ Bei der Binomialverteilung für ja-nein-Vorgänge handelt es sich um eine diskrete Verteilung mit ausschließlich ganzzahligen Ergebnissen ohne Zwischenwerte, so dass ein Balkendiagramm grundsätzlich die richtige Darstellungsform wäre. Durch die stetig prozentuale Wiedergabe wird zudem die starke Streuungsabhängigkeit von der Anzahl Daten (s. die Abb. im Abschnitt „Chi-Quadrat-Test“ unter Pkt. 4.3.1.3.) ausgeblendet. Weil zur

keitsmuster der Filialen „1“ und „3“ zeigen das erwartete symmetrische Verhalten und auf der X-Achse mit „ca. 8 %“ und „ca. 6 %“ den üblichen täglichen Anteil „Verzehr an Ort und Stelle“. Die Filialen „4“ und „5“ haben weitgehend glockenförmige Verteilungen mit **leichten Auffälligkeiten im Auslauf** bzw. Anstieg. Als eindeutig prüfungsrelevante Unternehmenseinheit sticht **Filiale „2“ mit einer unruhigen, zweigipfligen Häufigkeitskurve** hervor. Deshalb wird der Vertiefungsschwerpunkt gem. § 7 BpO⁶¹² auf die Filiale 2 gelegt.

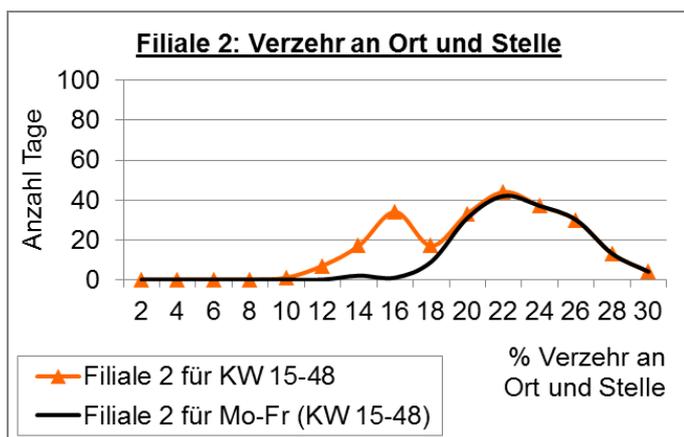
Durch die regelbasierte Zeitreihendarstellung zur Entwicklung der Anteile „Verzehr an Ort und Stelle“ pro Verkaufseinheit

sollen mögliche Ursachen für die auffällige Häufigkeitsverteilung von **Filiale 2** hinterfragt werden. Der anfänglich von ca. 15 % bis etwa zur 15. Kalenderwoche steigende Anteil mit anschließender Stabilisierung um 20 % legt den Gedanken nahe, dass dies das Häufigkeitsmuster beeinflusst haben könnte.



Als weitere Erkenntnis ist noch abzulesen, dass **Filiale 4** zwischen der 35. und 46. Kalenderwoche eine deutliche Absatzveränderung mit höherem Anteil offenbart. Das erklärt zwar den langsameren (rechtsschiefen) Auslauf der Häufigkeitsverteilung, wirft aus Prüfungssicht allerdings folgende Frage auf: Ist der Auslöser für den zeitweisen Effekt ein verändertes Verzehrverhalten der Kunden oder eine tlw. falsche Systembedienung?

Die weitere Prüfungsvertiefung zu Filiale 2 bzgl. der Trennung der Entgelte setzt auf der neuen Information auf, dass der typische Anteil „Verzehr an Ort und Stelle“ erst im Laufe des Jahres erreicht wird. Hierzu werden deren Daten auf den **Zeitraum „15. bis 48. Kalenderwoche“** beschränkt.⁶¹³ Da hierdurch aber die Doppelstruktur mit



zwei Gipfeln nicht einheitlich wird, liefert der anfängliche Anstieg des Anteils nicht die Erklärung. In einem zweiten Schritt wird den Tagesanteilen für den Betrachtungszeitraum „15. bis

Auswahl der besonders prüfungsrelevanten Filialen jedoch Abweichungen von der Verteilungssymmetrie am interessantesten sind, wurde diese übersichtlichere Darstellungsvariante gewählt.

⁶¹² § 7 S. 1 BpO: „Prüfungsgrundsätze: Die Außenprüfung ist auf das Wesentliche abzustellen.“

⁶¹³ Die Umstellung der Y-Achse von Prozentanteilen auf Tage dient der sich anschließenden Untersuchung nach Wochentagen. Die Y-Achsenkalierung ist dabei so gewählt, dass das Diagramm mit der Filialgegenüberstellung vergleichbar ist.

48. Kalenderwoche“ der Wochentag beige stellt. Durch das jeweilige Ein- und Ausblenden einzelner Wochentage unter ständiger Beobachtung des Häufigkeitsmusters lässt sich zügig erkennen, dass die Wochenendtage für den ersten Häufungshochpunkt mit niedrigerem Anteil verantwortlich sind. Der Ausschluss der Wochenendtage aus dem Häufigkeitsmuster löst die Doppelstruktur mit zwei Gipfeln folgerichtig auf. Somit hat sich durch systematische Verbildlichung einer unprüfbar großen Anzahl an Besteuerungsdaten eine weitere klare Prüfungsfrage herauskristallisiert: Warum weisen die Wochenendtage ausschließlich in Filiale 2 einen deutlich weniger Verzehr an Ort und Stelle als die anderen Wochentage auf? Kann es sich dabei um einen Bedienfehler des Wochenendpersonals handeln?

Auch dieses Beispiel zeigt, wie die gezielte Verknüpfung von regel- und musterbasierten Untersuchungen in großen Datenmengen eine leistungsstarke sowie zügige Risikoorientierung mit konkret herausgearbeiteten Prüfungsfragen erzielen kann.

5.3. Visualisierte Übereinstimmung (Schnittstellenprüfung)

Infolge der Digitalisierung von Unternehmensverwaltungen, durch welche „die Anzahl elektronisch überwachter Prozesse stetig [steigt]“⁶¹⁴, werden manuelle Datenschnittstellen zunehmend automatisiert. Dass gegenüber elektronischen Datenerfassungs-, -verarbeitungs- und -speicherprozessen nicht dieselbe professionelle Skepsis besteht wie gegenüber manuellen, stellt die Forderung nach „Tax-Compliance-Management-Systemen (TCSM)“ in der Literatur heraus: „Es müssen die Bereitschaft, die personelle Kapazität und das Budget vorhanden sein, damit die Softwarelösungen optimal in die Unternehmensprozesse integriert werden können, was insbesondere eine laufende Pflege der Programme und insbesondere der (unvermeidbaren) Schnittstellen erfordert.“⁶¹⁵ Genauso wie die GoBD mit dem Internen Kontrollsystem (IKS) „[...] Erfassungskontrollen (Fehlerhinweise, Plausibilitätsprüfungen), Abstimmungskontrollen bei der Dateneingabe, Verarbeitungskontrollen [und] Schutzmaßnahmen gegen die beabsichtigte und unbeabsichtigte Verfälschung von Programmen, Daten und Dokumenten“⁶¹⁶ als Voraussetzung für die Ordnungsmäßigkeit verlangt, muss zu einer zeitgemäßen **Formalprüfung** die Verifikation der Systemabläufe gehören.

Damit bspw. ein Warenwirtschaftssystem, wie es unterhalb ausschnittsweise als Schema abgebildet ist, die vom Handel gewünschten Funktionen erfüllen kann, müssen die benötig-

⁶¹⁴ Graf, „Anforderungen an die Vorlage von Einzelaufzeichnungen – Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen?“, BBK 18/2013, 875, 877.

⁶¹⁵ Liekenbrock, „Digital Tax Compliance – Praxishinweise zur technischen Umsetzung von Tax Compliance Management Systemen“, Ubg 2018, 43, 49.

⁶¹⁶ „Grundsätze zur ordnungsgemäßen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD)“, BMF-Schreiben v. 14.11.2014, BStBl. I 2014, 1450, Rz. 100.

ten Informationen über Stammdaten oder Eingaben beigesteuert werden und diverse Verarbeitungsprozesse automatisch ablaufen:



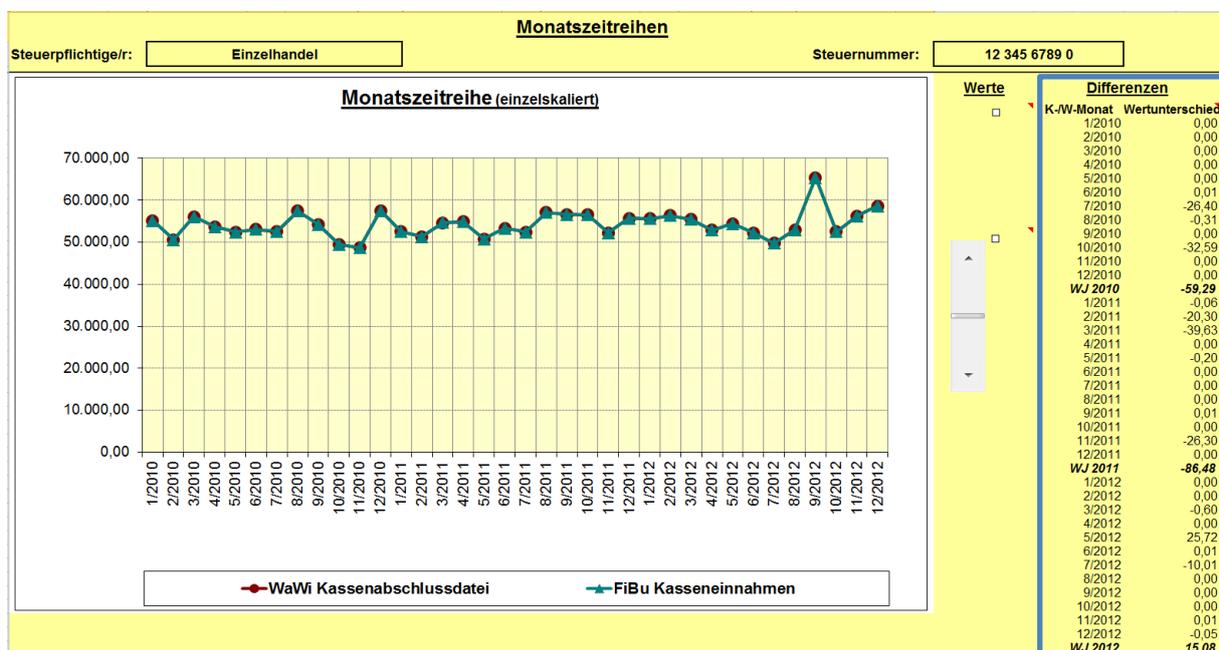
So versieht die **Kassenfunktion des Systems** die Einzelvorgänge bei der Erfassung mit allen Merkmalen zur protokollierenden Aufsummierung über den einzelnen Beleg bis zum täglichen Kassenabschluss. Dadurch wird die Abgrenzung für die Umsatzbesteuerung, die unterschiedlichen Zahlungswege etc. vorgenommen. Parallel werden die Daten der Lagerhaltung bzw. Inventur aktualisiert. Für die wertmäßige Inventursicht müssen zudem Bewertungen der Einzelpositionen über die Stammdaten oder manuell ergänzt werden. Alle verbildlichteten Abläufe können durch Falscheingaben oder fehlerhaft programmierte Prozesse zu unrichtigen Auswertungsergebnissen führen. Ziel der Schnittstellenprüfung⁶¹⁷ ist es, in möglichst effizienter Weise Fehler in den elektronischen Verwaltungsprozessen (weitest gehend) auszuschließen oder aufzudecken und einzugrenzen.

Da i.R. von Schnittstellenprüfungen üblicherweise (betriebliche) Fortlaufdaten mit einer chronologischen Abfolge geprüft werden sollen, bietet sich die Arbeitsweise mit **visualisierenden Zeitreihenanalysen** an. Im Gegensatz zur Schlüssigkeitsprüfung, bei der die Abhängigkeit zweier Datenreihen untersucht wird, ist hierbei allerdings eine möglichst vollständige Übereinstimmung zu erwarten. Deshalb müssen sich beide Datenreihen in der Zeitreihengrafik an derselben Y-Achse orientieren, weil jede automatische Größenanpassung einer Doppelskalierung das Ergebnis verzerren würde (s.a. den Vergleich der Zeitreihenvarianten im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)). Je größer die Gefahr erfassungsbedingter

⁶¹⁷ Vgl. z.B.: *Webel/Danielmeyer*, „Schnittstellenverprobung elektronischer Betriebsverwaltungen – das kommende Standardprüffeld“, StBp 2015, 353, sowie *Danielmeyer/Neubert/Unger*, „Praxiserfahrungen zu Vorsystemen – Prüffeld: Visualisierte Schnittstellenverprobung“, StBp 2016, 322.

Abweichungen einer Datenreihe gegenüber der anderen ist wie z.B. beim Vergleich von später oder aufsummiert erklärten Meldedaten mit der Buchführung, desto sinnvoller ist es, nach der SRP-Zeitreibensystematik „vom groben Überblick zu den Detailperspektiven“ vorzugehen (s. [Pkt. 4.2.2.](#)).

Im beispielhaften Fall eines Einzelhandelsbetriebs soll die Übereinstimmung der im Warenwirtschaftssystem erfassten Verkaufsvorgänge mit der Buchführung verifiziert werden. Dazu werden zunächst die täglichen Kassenabschlüsse aus dem Warenwirtschaftssystem den Buchführungsdaten gegenübergestellt. Weil es keinen Hinweis auf zeitliche Verschiebungen über die Monatsgrenzen hinaus gibt, wird direkt der einzelskalierte Zeitreihenvergleich vom Tabellenblatt „Monatszeitreihen“ in der SRP-Vorlage BWA (s. [Pkt. 4.2.5.2.](#)) betrachtet:

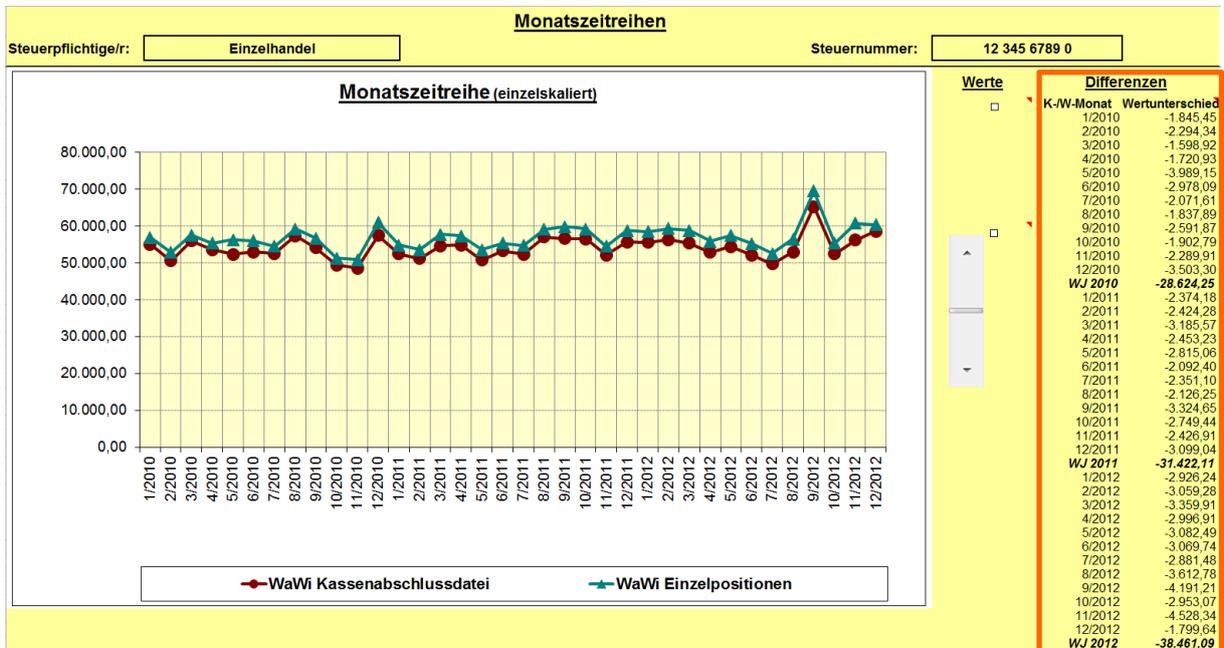


Die als papierenen Protokolle übergebenen Tageskassenabschlüsse stimmen mit den im Steuerbüro gebuchten Zahlen weitest gehend überein, wie die [tabellarische Ausgabe mit geringfügigen Monatsdifferenzen](#) bestätigt. Insoweit lässt sich umgehend feststellen, dass die manuelle Schnittstelle „Warenwirtschaftssystem des Betriebs ⇔ Buchführung im Steuerbüro“ ordnungsgemäß ist.

In einem zweiten Schritt werden die systeminternen Abläufe des Warenwirtschaftssystem gleichsam auf Schlüssigkeit hin überprüft. Hierzu werden die knapp 300.000 einzelnen Vorgangsdaten⁶¹⁸ des Warenwirtschaftssystem entsprechend der fortgeschrittenen Anwen-

⁶¹⁸ Sofern die Datenanzahl an Einzelvorgängen eines (Erst-)Erfassungssystems den möglichen Zeilenumfang von MS Excel® übersteigt, muss die tageweise Aufsummierung als Vorbereitungsschritt in IDEA® durchgeführt werden. Die Tagessummen können dann direkt in das Tabellenblatt „Vergleichsdaten“ der SRP-Vorlage BWA kopiert werden. Der Vorteil der verknüpfenden Arbeitsweise über das Tabellenblatt „eigene Notizen“ liegt in der Möglichkeit, die Zeitreihenanalysen nach weiteren Unterscheidungskriterien der Daten einzuzugrenzen (s. den Abschnitt „fortgeschrittene Anwendung der SRP-Vorlagen“ unter Pkt. 4.).

derung der SRP-Vorlagen auf dem Tabellenblatt „eigene Notizen“ abgelegt, tageweise per Pivot-Tabelle aufsummiert und auf das Tabellenblatt „Vergleichsdaten“ verknüpft (s. den Abschnitt „fortgeschrittene Anwendung der SRP-Vorlagen unter [Pkt. 4.](#)). Aus der Gegenüberstellung mit den täglichen Kassenabschlussdaten des Warenwirtschaftssystems resultieren dieses verbildlichte und tabellarische Ergebnis:



Während der Aufsummierung der Verkaufsvorgänge zu Tagessummen kommt es zu einer Wertminderung. Diese erstreckt sich ziemlich regelmäßig über den gesamten Prüfungszeitraum und bildet somit eine Systembesonderheit bzw. einen Systemfehler. Eine Überprüfung der für die Rechnungsstellung geschlüsselten Verkäufe ergibt, dass sie bereits in der Datenbank mit den Einzelvorgängen als „(noch) nicht werthaltig“ gelistet sind. Somit ist für einen sachverständigen Dritten kein Grund erkennbar, warum die erfassten Einzelverkäufe die Tagesabschlüsse übersteigen. Gerade systematische Manipulationen mit digitalen Hilfsmitteln [s. den Abschnitt „raffinierte (digitale) Manipulationswerkzeuge“ unter [Pkt. 2.1.](#)] können derartige Datenverfälschungen hervorrufen. Deshalb wirft die (Schnittstellen-)Prüfung des Betriebsverwaltungssystems in diesem Fall den Verdacht auf, dass die Zahlen der Buchführung nicht die korrekten Besteuerungsgrundlagen wiedergeben. Die Aufklärungsverantwortung trifft den Stpfl., weil er sich für diese schwierig verifizierbare Art der Aufzeichnung und Aufbewahrung entschieden hat.⁶¹⁹

Grundsätzlich ist eine Übereinstimmungsprüfung bspw. von Schnittstellen auch mit Hilfe von **musterbasierten Häufigkeitsanalysen** möglich. Bei dieser Herangehensweise überlagern sich zwei Häufigkeitsverteilungen mit evtl. Klassenabweichungen.

⁶¹⁹ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „Das Beweismaß [für eine Schätzung] hängt mithin davon ab, [...] wie weit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

Die Unterschiede der beiden gegenübergestellten Datenbestände lassen sich in Abhängigkeit von der Unterteilbarkeit des Klassifizierungsmerkmals (s. den Abschnitt „Auszählung von Informationen nach unterteilbaren Eigenschaften“ unter [Pkt. 4.3.](#)) ablesen wie z.B. die Erhöhung des Frauenanteils in der deutschen Gesellschaft durch die Gefallenen im 2. Weltkrieg gem. der Grafik zur „Altersstruktur Deutschlands nach Geschlecht 2011“ unter [Pkt. 5.](#) Insbesondere für viele ähnliche Daten wie bei Erlösen oder Kosten führt jedoch die zusätzliche Nutzung der chronologischen Information mittels verbildlichender Zeitreihenanalytik i.d.R. schneller zu konkreten Übereinstimmungsergebnissen.

5.4. Visualisierte Schlüssigkeit („sprechende Diagramme“)

Die korrekte Beurteilung, ob Vereinbarungen zwischen nahestehenden Personen fremdüblich und somit steuerlich anzuerkennen sind, setzt einen **guten Überblick über die tatsächlichen Verhältnisse** voraus. Dabei ist zum einen zu beachten, dass Dokumentationen wie bspw. zur Angemessenheit von Verrechnungspreisen eine subjektive (zusammenfassende) Aufzeichnung der realen Vorgänge darstellen. Ohne Einbeziehung der dazugehörigen Einzeldaten kann deshalb nicht von einer ausreichenden Sachverhaltsaufklärung ausgegangen werden. Zum anderen sind die Verhältnisse gerade bei Verrechnungen innerhalb internationaler Unternehmensverbände häufig so komplex gestaltet, dass sie alleine auf der Basis von Texten und Tabellen nicht vollständig erfasst werden können. Hinsichtlich der Beziehungen verbundener Unternehmen ist der Zusammenhang aus Verbildlichung und Verstehen bekannt, weshalb i.d.R. Organigramme erstellt werden. Die Daten zur Durchführung einer prüfungswürdigen Gestaltung zwischen nahestehenden Personen sind häufig weitaus komplexer.

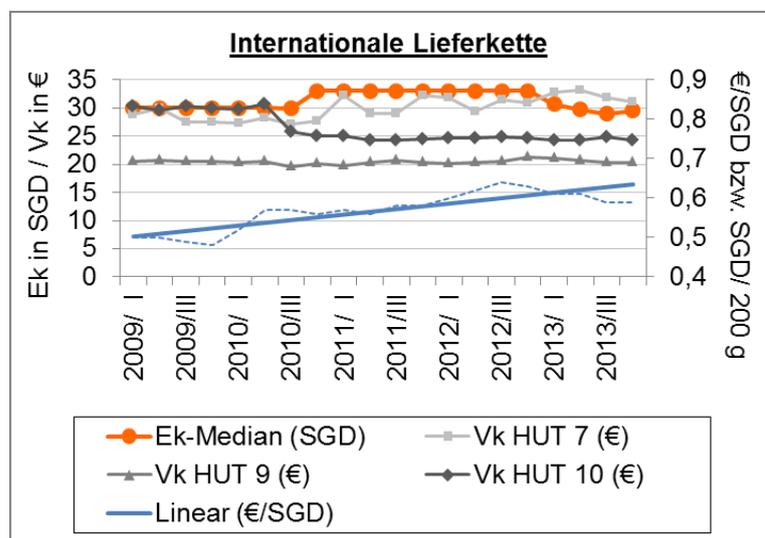
Die von der OECD vorgeschlagenen Maßnahmen zur Bekämpfung der illegalen Steuer-
vermeidung multinationaler Konzerne „BEPS“⁶²⁰ basieren u.a. auf einer transparenteren Offenlegung der Grundlagen zur internen Verrechnung. Dabei entstehen zwangsläufig noch mehr Daten zu Vergleichsmaßstäben bzw. kalkulatorischen Grundlagen, welche bei der betriebswirtschaftlichen Beurteilung der Fremdüblichkeit einbezogen werden sollen. Insofern ist die wahrnehmungsphysiologische Überlegenheit der Datenvisualisierung entscheidend (s. den Abschnitt „Visualisierung von Besteuerungsdaten“ unter [Pkt. 2.2.](#)).

Als Beispiel für eine visualisierende Sachverhaltsaufbereitung i.R. einer internationalen Verrechnungspreisprüfung mit Hilfe **systematischer Zeitreihenanalytik** dient die Bezie-

⁶²⁰ Internetseite der OECD „BEPS 2015 Final Reports: Final BEPS package for reform of the international tax system to tackle tax avoidance (BEPS = Base Erosion and Profit Shifting)“, nachzulesen unter: <http://www.oecd.org/ctp/beps-2015-final-reports.htm>.

hung eines deutschen Fertigungsunternehmens mit der Schwestergesellschaft in Singapur. In dem Unternehmensverbund (mit weiteren Gesellschaften in anderen Ländern) werden Millionen von Kunststoffkleinteilen hergestellt und weltweit vermarktet. Während die einfachen Produkte in Singapur gefertigt werden, erfolgt die Herstellung der aufwendigeren in den europäischen Fabriken. Innerhalb der Lieferkette „Produktion in Singapur ⇒ Lieferung nach Deutschland ⇒ Vermarktung weltweit“ ist die (beträchtliche) Roh-Marge zwischen Singapur und Deutschland sachgerecht aufzuteilen. In der Gesamtbetrachtung der Besteuerung des Unternehmensverbunds ergibt sich für die Eigentümer aufgrund der sehr unterschiedlichen Unternehmensbesteuerung beider Länder ein klares Motiv zur Gewinnverlagerung.⁶²¹ Deshalb soll eine Verbildlichung der chronologischen Preisentwicklungen und des wichtigen Einflussfaktors „Währungskurs“ möglichst viel Klarheit über den Beurteilungsgegenstand herstellen.

Im erstellten Diagramm⁶²² werden alle **Einkäufe der zehn Hauptumsatzträger** über den Prüfungszeitraum mit Quartals-Durchschnittspreisen in der Vertragswährung „SGD“ auf einen **repräsentativen⁶²³ Verlauf** zusammengefasst. Demgegenüber weisen die drei exemplarischen Verkaufspreise von Hauptumsatzträger „7“, „9“ und „10“ Entwicklungen unterschiedlicher Art auf.



Somit kommt als kalkulatorische Grundlage für die Einkaufspreisbildung nicht der jeweilige Absatzpreis der Produkte in Betracht. Im Zusammenhang mit dieser Erkenntnis verwundert die lang anhaltende ca. 10 %-ige Erhöhung der Einkaufspreise bei gleichzeitiger **Tendenz zur permanenten Währungsverteuerung**⁶²⁴. Unter fremden Geschäftspartnern wäre zu vermuten, dass das deutsche Unternehmen seine gro-

⁶²¹ Neue Züricher Zeitung online v. 14.2.2017: „Singapur geht den pragmatischen Weg: Im Stadtstaat erfolgt eine graduelle Anpassung der Unternehmensbesteuerung an international geltende Standards. Man will kein Paria sein“; nachzulesen unter: <https://www.nzz.ch/wirtschaft/unternehmensbesteuerung-singapur-geht-den-pragmatischen-weg-ld.145555>. Im Erlass zur „internationale[n] Einkunftsabgrenzung nach dem Maßstab des Fremdvergleichs“ v. 23.2.1983, BStBl. I 1983, 218, wird die besondere Situation gegenüber niedrig besteuerten Ländern explizit aufgegriffen: „1.5. Geschäftsbeziehungen zu Nahestehenden in niedrig besteuerten Gebieten“.

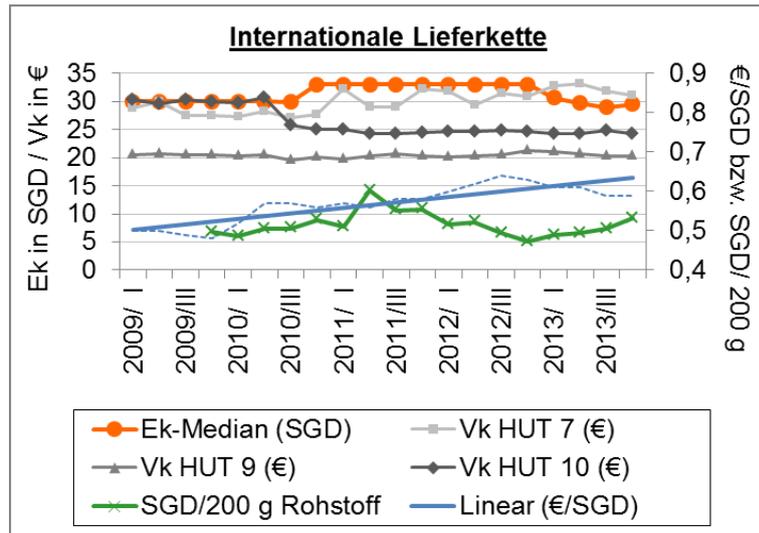
⁶²² Bildquelle (überarbeitet): Wolenski/Wähner, „Systematische Visualisierung von internationalen Gewinnverschiebungen“, IWB 3/2015, 105, 110.

⁶²³ Die Einzelpreise aller zehn Hauptumsatzträger folgen diesem Verlauf auf unterschiedlichem Preisniveau.

⁶²⁴ Über eine Anpassung der Y-Achsenkalierung wurde die Darstellung als Zwei-Ebenen-Diagramm erreicht, damit alle chronologischen Größenentwicklungen gut im Zusammenhang zu erkennen sind.

ßen Abnahmemengen nutzt, um den schwächer werdenden Euro über einen reduzierten Einkaufspreis tlw. auszugleichen. Gem. Unternehmensdaten verstärken sich die beiden Verteuerungseffekte jedoch mit der Folge einer Gewinnminderung in Deutschland.

Auf die offengelegten Verrechnungspreisverhältnisse gab die Wirtschaftsprüfungsgesellschaft als Grund einen „permanenten Preisanstieg des Rohstoffs“ an. Da zu dieser Begründung prüfbare Nachweise z.B. in Form auszugsweiser Rechnungen der Schwestergesellschaft aus Singapur fehlten, sollte eine Erweiterung der Sachverhaltsverbildlichung das Ar-



gument hinterfragen. Dazu wurden Rohstoffkurse des betreffenden Kunststoffes von einer Internetseite heruntergeladen und in das Diagramm einbezogen.⁶²⁵ Trotzdem hierbei für die ersten drei Quartale keine Daten erhältlich waren, erwies sich der Einwand als nicht tragfähig: Eine relevante⁶²⁶ Verteuerung des Rohstoffs erfolgte danach erst

im II. Quartal 2011 und somit ein halbes Jahr nach der vertraglichen Erhöhung aller Einkaufspreise. Außerdem fiel der Rohstoffpreis innerhalb eines Jahres wieder auf sein altes Niveau zurück, während die Einkaufspreiserhöhung noch deutlich länger Bestand hatte. Infolge dieser effektiven und effizienten Aufbereitung einer prüfungsbedürftigen Vereinbarung zwischen nahestehenden Personen trifft nun den Stpfl. die Darlegungsverantwortung, dass die festgelegten Verrechnungspreise dennoch fremdüblich und deshalb steuerlich anzuerkennen sind⁶²⁷.

Immer wenn eine Schlüssigkeitsfrage Bezug zu merkmalsbezogenen Häufigkeitsverteilungen hat, kann eine **visualisierende Musteranalyse** als Entscheidungsgrundlage mehr Klarheit für den zu beurteilenden Sachverhalt erreichen wie insbesondere in den folgenden Konstellationen:

⁶²⁵ Die offiziellen Rohstoffpreise wurden auf die Bezugsmenge „200 g“ umgerechnet, damit die Sekundärachse die Preisentwicklung gut nachvollziehbar im Diagramm platziert.

⁶²⁶ Sowohl die Währungskurs- als auch die Rohstoffpreisentwicklung werden im Diagramm dadurch betont, dass die sekundäre Y-Achse bei einem Wert oberhalb von Null startet.

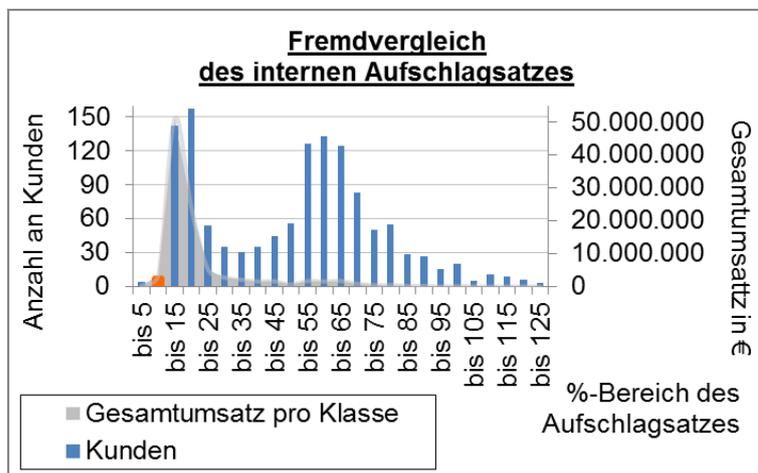
⁶²⁷ Tipke/Kruse, AO/FGO, 150. Lieferung 10/2017, § 158 AO, Rz. 9: „Das Beweismaß [für eine Schätzung] hängt mithin davon ab, wie viel und in welcher Form der Stpfl. Aus seiner Verantwortungssphäre zur Sachaufklärung beiträgt und wieweit die FinBeh. auf dessen Mitwirkung im konkreten Einzelfall angewiesen ist.“

- ✓ Ein (einfacher) Sachverhalt soll anhand einer Vielzahl von Vergleichsvorgängen beurteilt werden (s. das nachfolgende Beispiel).
- ✓ Ein komplexer Sachverhalt mit uneinheitlicher Beurteilungsgröße soll durch Gegenüberstellung auf Fremdüblichkeit geprüft werden (s. das nachfolgende Beispiel).
- ✓ Für einen merkmalsbezogen klassifizierbaren (s. den Abschnitt „Auszählung von Informationen nach unterteilbaren Eigenschaften“ unter [Pkt. 4.3.](#)) Sachverhalt besteht eine Erwartung als Beurteilungsmaßstab wie bspw. bei der Verpflichtung zur gleichmäßigen Einzelbewertung von Forderungen unabhängig vom Ausgangswert⁶²⁸.

Im Beispielsfall soll auf der Grundlage mehrerer tausend elektronischer Rechnungsausgangsdaten die Fremdüblichkeit der Verrechnung eines deutschen Unternehmens an ein verbundenes in Spanien beurteilt werden. Die kalkulatorische Gegenüberstellung der vorgangsbezogenen Informationen „Umsatz“ und „Kosten“ in der Datenbank ergibt eine erhebliche **Margendifferenz mit 8,1 % intern** gegenüber 22,3 % extern.

Durch die nebenstehende Häufigkeitsauswertung⁶²⁹ der kundenbezogenen Aufschlagsätze wird jedoch die Ungleichmäßigkeit der Kalkulation ersichtlich.

Mit den beiden Konzentrationsbereichen „> 10 % bis 20 %“ sowie „> 50 % bis 65 %“ müssen



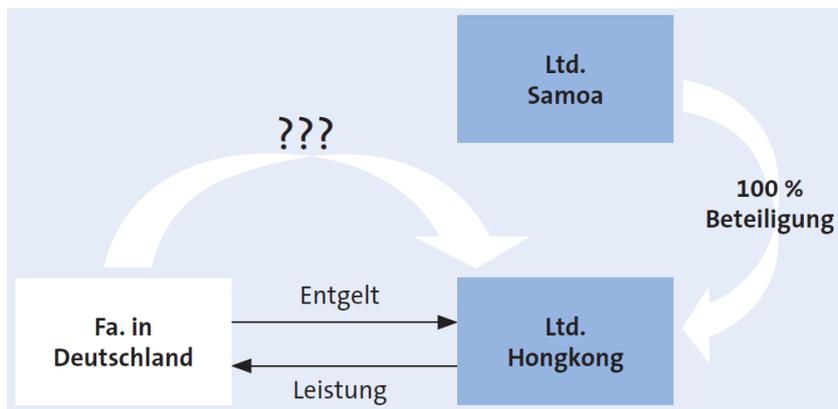
offensichtlich zwei Unternehmensbereiche differenziert werden. Die zusätzliche Abb. der Rechnungsvolumina pro Aufschlagsatzklasse verdeutlicht zudem, dass die höheren Margen zwar viele Kunden mit allerdings geringen Rechnungsbeträgen betroffen haben. Da sich durch einen gezielten Vergleich der beiden Vorgangsarten herausstellt, dass es sich bei den Sachverhalten mit niedrigerer Marge um Warenlieferungen und bei denen mit höherer um Ersatzteile handelt, muss der Fremdvergleich diesbezüglich konkretisiert werden. Mit Hilfe weiterer Datenfilter zu kalkulationsrelevanten Faktoren wie bspw. den verkauften Produktgruppen kann der verbildlichte Fremdvergleich zügig auf eine sachgerechte Vergleichsgruppe eingegrenzt werden.

⁶²⁸ Vgl. *Wolenski/Weiskopf*, „Praxisfall: Einsatz indirekter Prüfungsmethoden in der Großbetriebsprüfung am Beispiel der Einzelwertberichtigung“, BBK 3/2013, 128.

⁶²⁹ Bildquelle (überarbeitet): *Werner/Wähner*, „Verbildlichtende Prüfung am Beispiel der Gewinnverlagerung – Die Überlegenheit der Visualisierung von Datenregeln und -verteilungen gegenüber Kennzahlen“, IStR 19/2016, 811.

Neben einer verbesserten Sachverhaltsermittlung dient eine visualisierte Schlüssigkeitsprüfung auch als klarere Verhandlungsgrundlage in Prüfungsbesprechungen („**sprechende Diagramme**“). Es hat sich inzwischen in vielen Prüfungen gezeigt, dass fokussierte Verbildlichungen der Daten, die einen auf Fremdüblichkeit zu beurteilenden Sachverhalt betreffen, zu Ergebnissen führen, welche weniger vom Verhandlungsgeschick der Parteien aber dafür mehr von der Gesetzmäßigkeit der Abläufe geprägt sind.

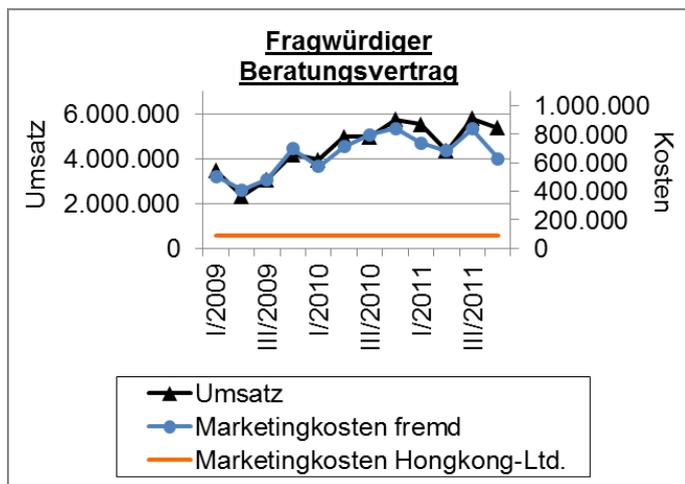
Als Beispiel für den grundsätzlichen Nutzen der Visualisierung von prüfungswürdigen



Gestaltungen dient dieses einfache Konstrukt. „Fraglich ist, ob Beteiligungen der Gesellschafter der deutschen D-GmbH an der Ltd. in Hongkong über eine Ltd. in Samoa indirekt gehalten werden. Samoa erteilt an deutsche Finanz-

behörden zumindest nur sehr schleppend Auskünfte, so dass der Standort bewusst gewählt sein könnte, um die Unternehmensverbindung der D-GmbH und der Ltd. Hongkong im Verborgenen zu lassen. Auf diese Weise können über einen Leistungsvertrag mit fingierten oder überbewerteten Betriebsausgaben Gewinne aus Deutschland in einen Staat mit erheblich niedrigerer Unternehmensteuerquote verschoben werden.“⁶³⁰

Zu den **pauschalen Marketingkosten der Ltd. aus Hongkong** können keine überprüfbareren Tätigkeitsnachweise vorgelegt werden. Um ihre fehlende Fremdüblichkeit herauszustellen,



wird deren Entwicklung und Abhängigkeit vom Erfolg der D-GmbH im abgebildeten Zeitreihendiagramm **vergleichbaren Beratungskosten** gegenübergestellt. Unter der angemessenen groben Vergleichsperiode „Quartal“ ist ein eindeutiger Zusammenhang zwischen dem zunehmenden Erfolg der D-GmbH als Beratungsadressat und den berechneten Kosten zu erkennen.

Selbst die zwischenzeitlichen Umsatzenschwankungen werden in der Anpassung der Marke-

⁶³⁰ Wolenski/Wähnert, „Systematische Visualisierung von internationalen Gewinnverschiebungen“, IWB 3/2015, 105, 106 mit Text u. Bild.

tingkosten beantwortet. Demgegenüber bleiben die Aufwendungen für die Ltd. in Hongkong unverändert. Für den Fortgang der Prüfung bedeutet das visualisierte Ergebnis, dass die Eigentumsverhältnisse der Ltd. aus Samoa unbedingt aufzuklären sind. Diese kommunizierte Erkenntnis hat im „Beispielsfall [...] zur Offenlegung [...] geführt, [wonach] die Gesellschafter der D-GmbH über die Ltd. in Samoa auch die Anteile an der Hongkong-Ltd. hielten und sich somit der Verdacht nahestehender Unternehmen bestätigte.“⁶³¹ Demnach hat die unmissverständliche Aufbereitung der Sachverhaltsdaten mindestens eine deutliche Beschleunigung der ausreichenden Aufklärung für die sachgerechte steuerliche Beurteilung erzielt.

5.5. Ausblick zur visualisierenden Datenanalyse

Die **Entwicklung der (professionellen) Datenanalyse** folgt offenbar diesen zwei Trends:

1. weitest mögliche **Automatisierung** durch „künstliche Intelligenz (KI)“ auf der Basis selbstlernender Algorithmen⁶³² sowie
2. **quantitatives** Ausreizen des inzwischen akzeptierten Visualisierungsvorteils durch die gleichzeitige Bereitstellung vieler Datenverbildlichungen⁶³³.

Aus Angebotssicht sind diese Strategien nachvollziehbar, weil sie nur eine geringe Auseinandersetzung mit den AnwenderInnen erfordern. Auf der einen Seite erhält das Prüfungspersonal ein fertiges Datenanalyseprodukt und auf der anderen werden die – nach aktueller Einschätzung – wichtigsten Datenverbildlichungen komplett zur Verfügung gestellt. In der Konsequenz können die schwieriger (gleichmäßig) zu adressierenden Voraussetzungen für eine qualifizierte Datenanalyse „Kenntnis der Methodengrundlagen“ und „technische Fähigkeiten zur Umsetzung“ auf das Notwendigste beschränkt werden.

Nach Ansicht des Autors werden beide Strategien aus den folgenden Gründen **keine optimale Breitenanwendung der digitalen Datenanalyse** erreichen können:

- ! Die Informationsauswertung im menschlichen Gehirn erfolgt chaotisch-assoziativ⁶³⁴ und damit auf einem anderen Weg als die – vom Menschen erschaffene – hochgradig ge-

⁶³¹ Wolenski/Wähnert, „Systematische Visualisierung von internationalen Gewinnverschiebungen“, IWB 3/2015, 105, 109.

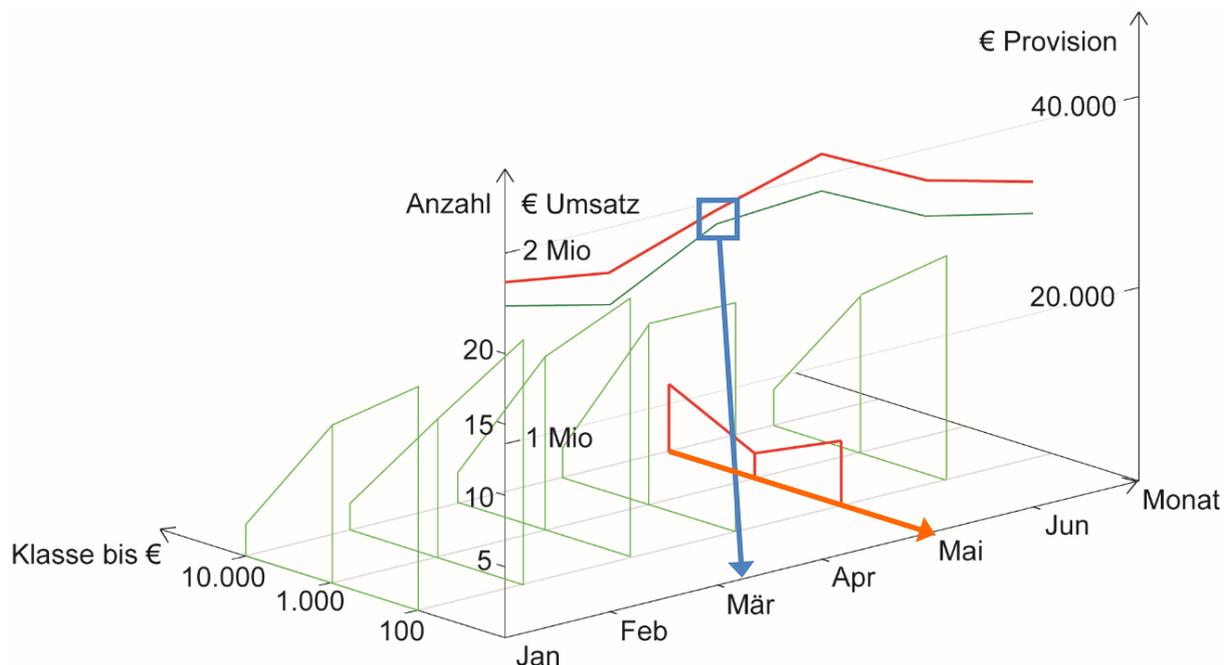
⁶³² Kompenhans/Wermelt, „Big Data – Potenziale für innovative Abschlussprüfungen“, BB 2018, 299: „Die an Systemschnittstellen erforderlichen manuellen Tätigkeiten werden immer mehr durch den Einsatz von Robotic Process Automation ersetzt, Entscheidungsprozesse werden dabei durch den Einsatz programmierter, selbstlernender Entscheidungsheuristiken, die auf Basis künstlicher Intelligenz arbeiten, unterstützt.“

⁶³³ Vgl. z.B.: Internetblog der Fa. Audicon zu IDEA V10: „Grafische Darstellungen lassen sich im Allgemeinen leichter erfassen als die Darstellung der Daten in Tabellenform. Daher ist ein besonderes Feature des neuen Releases die Visualisierung“, nachzulesen unter: <https://blog.audicon.net/audit-analytics/idea10-visualisierung-und-ermitteln.html> oder Visualisierungsprogramm „Visual Analytics“ des amerikanischen, weltweit operierendes Softwarehauses Statistical Analysis Systems, nachzulesen unter: http://www.sas.com/de_de/software/business-intelligence/visual-analytics.html.

ordnete Datenverarbeitung. Wird das Prüfungspersonal wie Konsumenten an das Ende der Datenanalyseprozesse gestellt, verzichtet die Sachverhaltsaufklärung (weitgehend) auf die **menschlichen Kompetenzen „Intuition“ und „Erfahrung“**. Eine optimale Sachverhaltsaufklärung kann nur durch die bestmögliche Kombination aus Personalkompetenzen und Datenanalyse gelingen.

- ! (Elektronische) Datenanalysen ergeben nicht automatisch richtige Ergebnisse. Wie das Beispiel unterschiedlicher Zeitreihenvarianten als Grundlage für die Schlüssigkeitsbeurteilung derselben Betriebsdaten *im Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)* unmissverständlich aufzeigt, können sie zu unzutreffenden Schlussfolgerungen verleiten. Deshalb ist ein **ausreichendes Verständnis der vielen Datenanalysemethoden** unverzichtbar, um geeignete Datenanalysemittel für eine Prüfungsfrage zu wählen.

Die heutigen Möglichkeiten der digitalen Datenverarbeitung sollten im Prüfungsbereich dazu genutzt werden, **wichtige (visualisierende) Prüfungsansätze sinnvoll und variabel miteinander zu verknüpfen**. So werden in der abgebildeten Zeichnung⁶³⁵ regel- und musterbasierte Verbildlichungen zur Untersuchung von umsatzabhängigen Provisionen in einer Analyse kombiniert:



Die Frontalebene aus der X- und den beiden Y-Achsen verfolgt die Abhängigkeit beider Betriebsgrößen als doppelskalierter Zeitreihenvergleich. Mit dem hierfür besonders **geeigneten (zur unterschiedlichen Eignung bestimmter Zeitreihenvarianten für die Plausibilitätsprüfung s.**

⁶³⁴ Intuitiv scheint dieser Unterschied den Führungskräften der Wirtschaft bewusst zu sein: „Manager vertrauen mehr auf ihr Bauchgefühl als auf B[usiness]I[n]telligence“ zitiert CRN online v. 24.2.2010 eine entsprechende Studie.

⁶³⁵ Bildquelle: *Wähnert*, „Das große Potenzial von Strukturvergleichen in der Revision“, StC 4/2012, 18, 23.

den Abschnitt „Methodenkenntnis“ unter [Pkt. 4.2.](#)) Zeitreihenmodell ist umgehend zu erkennen, dass der Monatswert an Provisionen im März verhältnismäßig größer ausfällt als in allen anderen Monaten. Die zusätzliche Z-Achse zerlegt i.V. mit der primären Y-Achse die monatlichen Provisionssummen in die Anzahl an Einzelvorgänge pro logarithmische⁶³⁶ Größenklasse. Auf diese Weise wird neben der Höhe an monatlichen Provisionskosten auch ersichtlich, ob diese anders abgerechnet wurden. Entsprechend legt die Kombi-Analyse offen, dass im Monat „Mai“ doppelt so viele (betriebsbezogen) sehr große Provisionsrechnungen und deutlich weniger mittlere und kleine vorliegen. Dank der gleichzeitigen Anzeige der regel- und musterbasierten Auffälligkeiten lässt sich besser beurteilen, welche prüfungsrelevanter sind. Außerdem haben derartige Regel- und Musterstörungen häufig einen inhaltlichen Zusammenhang (s. den Abschnitt „kombinierte regel- und musterbasierte Datenverbildlichung“ unter [Pkt. 5.2.](#)), wodurch resultierende Prüfungsfragen konkretisiert werden.

Das Lesen derartiger dreidimensionaler Untersuchungen ist für datenanalytisch geschultes Prüfungspersonal problemlos erlernbar. Durch die Möglichkeit der freien Zuordnung beider Datenreihen und der Klassifizierungsmerkmale (s. den Abschnitt „Ausählung von Informationen nach unterteilbaren Eigenschaften“ unter [Pkt. 4.3.](#)) entstünde ein extrem leistungsfähiges Prüfungswerkzeug.

⁶³⁶ Eine logarithmische Klassenbildung belegt die gleichmäßigen Intervalle mit linear ansteigenden Exponenten wie in der Beispielsgrafik mit „ $10^2 - 10^3 - 10^4$ “ (entsprechend „100 – 1.000 – 10.000“) und lenkt dadurch den Fokus auf die höherwertigen Rechnungen.

6. (Ausgewählte) Literatur zur „Neuen Prüfungstechnik“⁶³⁷

-  *Achilles*, „Kassen-Nachschaue nach § 146b AO“, DB 1-2/2018, 18.
-  *Achilles*, „Kassenführung – Bargeschäfte sicher dokumentieren“, DATEV-Verlag, 2. Auflage 2018.
-  *Achilles*, „Kassenführung in bargeldintensiven Unternehmen“, BoD Verlag, 2014.
-  *Anders/Gärtner*, „Manipulationssicherheit in der Bargeldbranche durch Einzelaufzeichnungen“, Stbg 2/2016, 49.
-  *Anders/Rühmann*, „Aufbewahrungspflicht für Protokolle zur Kassenprogrammierung – Nichtaufbewahrung als schwerer Ordnungsmangel der Kassenführung“, BBK 13/2013, 627.
-  *Barthel*, „Skalierte Außenprüfung“, Stbg 3/2018, 115.
-  *Barthel*, „Betriebsprüfung: Die griffweise Schätzung“, Stbg 7-8/2017, 315.
-  *Baselt*, „Einführung in die Strukturanalyse“, Lehrgangsskript der Bundesfinanzakademie.
-  *Becker*, „Außenprüfung digital – Prüfungsmethoden im Fokus“, DStR 24-25/2016, 1386 u. 1430.
-  *Becker*, „Verprobung der Pfandverbindlichkeiten von Mineralbrunnen mit Hilfe der Visualisierung“, Bonner Bp-Nachrichten 12/2015, 19.
-  *Becker*, „Textbausteinsystem Schätzung“, juris (finanzverwaltungsintern): Fachinfosystem Bp/Allgemeine Prüfungshinweise/Arbeitspapiere und Leitfäden/TBS Schätzung.
-  *Becker/Danielmeyer/Neubert/Unger*, „‘Digitale Offensive‘ der Finanzverwaltung: Die Schnittstellen-Verprobung (SSV)“, DStR 51/2016, 2983.
-  *Becker/Giezek/Webel/Wähnert*, „Der Beanstandungsanlass nach § 158 AO“, DStR 32/2016, 1878.
-  *Becker/Schumann/Wähnert*, „Aktuelle Rechtsprechung zur Neuen Prüfungstechnik: Summarische Risikoprüfung (SRP) und Quantilsschätzung“, DStR 23/2017, 1243.
-  *Becker/Wiethölter*, „Aufzeichnungspflichten beim Einnahmen-Überschuss-Rechner gemäß § 4 Abs. 3 EStG – Der falsch verstandene Beschluss des BFH vom 16. 2. 2006 X B 57/05“, StBp 2009, 377.
-  *Becker/Wiethölter*, „Aufbewahrung der Belege und Einzelaufzeichnung von Betriebseinnahmen bei Gewinnermittlern nach § 4 Abs. 3 EStG - oder: die Ruhe nach dem durch Beschluss des BFH vom 16.2.2006 verursachten Sturm“, StBp 2009, 239.

⁶³⁷ Grau gedruckte Quellen sind nach Kenntnisstand des Autors nur innerhalb der Finanzverwaltung erhältlich.

-  *Bendisch*, „Die Strukturanalyse als Instrument der Betriebsprüfung“, wissenschaftliches Gutachten im Auftrag des Bundesministeriums für Finanzen, Mai 1980 (Az. BFA – P 3220 – 5/80).
-  *Bundesfinanzakademie*, „Podiumsdiskussion: Monetary Unit Sampling (MUS)“ sowie „Zehnjähriges Jubiläum der Summarischen Risikoprüfung“, Quartals-Info I/2018, downloadbar unter: https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Ministerium/Geschaeftsbereich/Bundesfinanzakademie/Aktuelles/BFA_Quartalsinfo_2018_I.html.
-  *Bundesfinanzakademie*, „Podiumsveranstaltung ‚Prüfung und Verprobung von Big Data‘“, QuartalsInfo I/2015, 11, downloadbar unter: https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Ministerium/Geschaeftsbereich/Bundesfinanzakademie/Aktuelles/BFA_QuartalsInfo_2015_I.html.
-  *Brähler/Bensmann/Emke*, „Der Einsatz mathematisch-statistischer Methoden in der digitalen Betriebsprüfung, Ilmenauer Schriften zur Betriebswirtschaftslehre 4/2010, Verlag proWiWi e.V., 2010.
-  *Brinkmann*, „Die Begründung des Sicherheitszuschlags“, StBp 2018, 212.
-  *Brinkmann*, „Schätzungen im Steuerrecht“, Erich Schmidt Verlag, 4. Auflage 2017.
-  *Brinkmann*, „Der so genannte Sicherheitszuschlag“, StBp 2014, 29 u. 69.
-  *Brinkmann*, „Strafrechtliche Aspekte der Schätzung“, StBp 2013, 250 u. 291.
-  *Danielmeyer/Neubert/Unger*, „Die Betriebsprüfung von A–Z“, StBp 2017, 291.
-  *Danielmeyer/Neubert/Unger*, „Praxiserfahrungen zu Vorsystemen-Prüffeld – Visualisierte Schnittstellenverprobung“, StBp 2016, 322.
-  *Danielmeyer/Neubert/Unger*, „Schnittstellenverprobung im Rahmen der SRP“, Bonner Bp-Nachrichten 8/2016, 25.
-  *Danielmeyer/Neubert/Unger*, „Schnittstellen-Verprobung – Das Prüffeld mit Zukunft!“, Bonner Bp-Nachrichten 4/2016, 17.
-  *Diller/Schmid/Späth/Kühne*, „Zifferntests in der Betriebsprüfung – Chancen und Risiken“, DStR 6/2015, 311.
-  *Diverse*, „Kassenprüfungen und Zeitreihenvergleich in der Praxis“, BBK-Sonderheft 2016.
-  *Ehlers*, „Tax Compliance – Neue Beratungsfelder bei Betreuung kleiner und mittlerer Unternehmen“, NWB 18/2012, 1535.
-  *Ettinger/Papenfuß*, „Visualisierung und Verrechnungspreise oder ‚ein Bild sagt mehr als tausend Worte‘“, Bonner Bp-Nachrichten II/2018, 26.
-  *Franke-Roericht*, „‚Wenn die Kasse klüngelt‘: Beihilfe durch Verkauf von Kassenmanipulationssoftware“, PStR 3/2015, 66.

-  *Gebbers*, „Durchführung eines z-Tests zur Prüfung der Umsätze“, StBp 2012, 308, 340 u. StBp 2013, 13.
-  *Gebbers*, „Analyse der Rechtsprechung zum Chi-Quadrat-Test in der Außenprüfung - Anwendung zur Prüfung von Erlösen und Fahrtenbüchern“, StBp 2008, 209 u. 290.
-  *Gebbers*, „Die Strukturanalyse“, StBp 2007, 225, 257 u. 289.
-  *Giezek*, „Monetary Unit Sampling“, Gabler Verlag, 2011.
-  *Giezek/Rupprecht/Wähnert*, „Wie sicher muss ein Verprobungsergebnis sein? – Anleihen aus der Kriminalistik zur sinnvollen Einordnung des Beanstandungsanlasses“, BBK 5/2017, 236.
-  *Giezek/Wähnert*, „Wahrscheinlichkeitstheorie in der Betriebsprüfung – Von der qualifizierten Stichprobenprüfung zur Ziffernanalyse“, DB 9/2018, 470.
-  *Giezek/Wähnert*, „Schätzungsmethoden der Betriebsprüfung im Vergleich – Einordnung der wichtigsten Schätzungsverfahren durch Praxis und Wissenschaft“, BBK 21/2017, 998.
-  *Giezek/Wähnert/Becker*, „Anwendung des Monetary Unit Sampling (MUS) in der steuerlichen Betriebsprüfung – Qualifizierte Untersuchung und Bewertung von Stichproben“, StBp 12/2016, 347.
-  *Graf*, „Anforderungen an die Vorlage von Einzelaufzeichnungen - Ende der ‚Schonzeit‘ für Daten aus Vor-Systemen“, BBK 18/2013, 875.
-  *Graf*, „Paradigmenwechsel durch die Datenanalyse – Datenanalyse und Mitwirkungsverpflichtung bei steuerlichen Betriebsprüfungen – eine Symbiose?“, BBK 7/2012, 331.
-  *Härtl*, „Einzelaufzeichnungspflicht und Datenzugriff auf digital aufgezeichnete Geschäftsvorfälle von Bargeschäften – (Kein) Datenzugriff bei Registrierkassen, Apotheken?“, StBp 2014, 1 u. 45.
-  *Härtl*, „Der Datenzugriff bei Überschussrechtern nach § 4 Abs. 3 EStG im Licht des BFH-Urteils vom 24.6.2009 (VIII R 80/06), im Besonderen auf gesetzlich nicht vorgeschriebene digitale Aufzeichnungen im Sinne des § 146 Abs. 6 AO“, StBp 2009, 337.
-  *Härtl/Schieder*, „Ordnungsmäßigkeit digital geführter Erlösaufzeichnungen – Elektronische Registrierkassen und digitale Erlöserfassungssysteme im Brennpunkt des Steuerrisikos Erlöskürzung“, StBp 2011, 33, 68 u. 97.
-  *Harle/Olles*, „Die moderne Betriebsprüfung“, NWB-Verlag, 3. Auflage 2017.
-  *Hauschildt/Wähnert*, „Vollständigkeitsprüfung mit IDEA – Was bedeutet die elektronische Betriebsprüfung für Unternehmer und Berater“, NWB 40/2007, 3539.
-  *Henn*, „Verfahrensdokumentationen nach den GoBD“, DB 5/2016, 254.
-  *Huber*, „Die Zukunft des steuerlichen Risikomanagements im Einnahmebereich“ – Der Vertrauensvorschuss nach § 158 AO im historischen Zusammenhang und im Umfeld der Aufzeichnungstechnik der Gegenwart“, Protokoll zum 130. Bochumer Steuerseminar für Praktiker und

Doktoranden vom 14.11.2014 (Wilms), downloadbar unter: http://www.fachanwalt-fuer-steuerrecht.de/pdf/Protokoll_20100091.pdf.

-  *Huber*, „Die Verwertung von Echtdateien aus Kassennachschau mit Datenzugriff – Praxisbeispiele zur interaktiven Prüfungstechnik bei Bargeschäften“, BBK 16/2014, 757.
-  *Huber*, „Gedanken zur Zukunft des steuerlichen Risikomanagements im Erlösbereich“, StBp 2014, 121, 153, 185, 221, 245, 277 u. 317.
-  *Huber*, „Praktische Erfahrungen mit der Kassenführung in Österreich – Umsetzung der Kassenrichtlinie 2012 und Kassennachschau durch die Finanzpolizei“, BBK 6/2014, 286.
-  *Huber*, „Steueraufsicht und Betriebsprüfung in der Zeit der Kassenandroiden und ohne INSIKA – Grundgedanken, Ziele, Risiken und zweitbeste Lösungen“, StBp 2012, 301, 333, StBp 2013, 5, 35, 71 u. 93.
-  *Huber*, „Risikomanagement in der Außenprüfung“, Protokoll zum Bochumer Steuerseminar für Praktiker und Doktoranden vom 19. Februar 2010 (Gabert), downloadbar unter: http://www.fachanwalt-fuer-steuerrecht.de/pdf/protokoll_20100002.pdf.
-  *Huber*, „Die neue Prüfungstechnik – Mathematische Weiterentwicklungen – Theorie der Prüfungsmethoden & praktische Revisionsbeispiele“, LexisNexis-Verlag ARD Orac, 2009.
-  *Huber*, „Über Registrierkassen, Phantomware, Zapping und Fiskallösungen aus Deutschland und Österreich“, StBp 2009, 153, 185, 217, 253, 286, 317 u. 342.
-  *Huber*, „Digitale Ziffernanalyse versus Strukturanalyse und die logische Herleitung von Benford (NBL)“, StBp 2008, 241 u. 273.
-  *Huber*, „Weiterentwickelte und neue Methoden der Überprüfung, Verprobung und Schätzung“, StBp 2002, 199, 233, 258 u. 293.
-  *Huber/Reckendorf/Zisky*, „Die Unveränderbarkeit der (Kassen-)Buchführung nach § 146 Abs. 4 AO im EDV-Zeitalter und INSIKA – Teil 1: Historische Wurzeln und Realzustand, Teil 2: Probleme, Ursachen und Lösungen, Teil 3: Konzept, Technik, Praxis“, BBK 12/2013, 567, BBK 13/2013, 610 u. BBK 14/2013, 663.
-  *Huber/Seer*, „Steuerverwaltung im 21. Jahrhundert: Risikomanagement und Compliance“, StuW 4/2007, 355.
-  *Huber/Wähnert*, „Der systematische Einsatz neuer Prüfverfahren in der Finanzverwaltung – Neue interaktive Prüfungstechnik (NiPt)“, NWB 36/2009, 2814.
-  *Huber/Wähnert*, „Das Kölner Zeitreihenurteil und das Projekt ‚NiPt‘“, StBp 2009, 207.
-  *Kläne*, „Registrierkassen im Fokus der Außenprüfung – Aktuelle Entwicklung und Rahmenbedingungen“, NWB 13/2013, 923.
-  *Klingebiel*, „Ordnungsgemäße Kassenführung und sachliche Richtigkeit der Buchführung – Voraussetzungen für Hinzuschätzungsberechtigungen bei Außenprüfungen“, NWB 41/2008, 3849.

- ☞ *Knief*, „Eine BWA-Mindestanalyse - Eine logische Konsequenz der Digitalisierung des Rechnungswesens und der verschiedenen gesetzlichen Anforderungen“, StBp 1/2017, 21.
- ☞ *Knief*, „Die grafische Betrachtung der unterjährigen Unternehmensentwicklung für mittelständische Unternehmen – wider die Stichtagsanalyse“, DStR 43/2018, 2080.
- ☞ *Kompenhans/Wermelt*, „‘Big Data’ – Potenziale für innovative Abschlussprüfungen“, BB 2018, 299.
- ☞ *Krumm*, „Rechtsfragen der digitalen Betriebsprüfung – (Summarische Risikoprüfung)“, DB 20/2017, 1105.
- ☞ *Kulosa*, „Herausforderungen der digitalen Betriebsprüfung, u.a. bei bargeldintensiven Betrieben“, FR 2017, 501.
- ☞ *Kulosa*, „Replik zum Beitrag von Wolenski zum Zeitreihenurteil des X. Senats des BFH“, Stbg 6/2016, 271.
- ☞ *Liekenbrock*, „Digital Tax Compliance – Praxishinweise zur technischen Umsetzung von Tax Compliance Management Systemen“, Ubg 2018, 43.
- ☞ *Mack*, „Überprüfung und Schätzung im so genannten bargeldintensiven Handel – Problematik der Prüfsysteme von elektronischen Kassensystemen“, AO-StB 1/2016, 17.
- ☞ *Madauß*, „Statistische-mathematische Methoden in Besteuerungs- und Steuerstrafverfahren“, NZWiSt 1/2014, 24.
- ☞ *Mehret/Wähnert*, „Prüfungsnetze vs. Einzelmethoden: Ein wichtiger, (bisher) verkannter Vorteil der Summarischen Risikoprüfung (SRP)“, DStR 6/2018, 314.
- ☞ *Neubert*, „Die digitalen Methoden der modernen Betriebsprüfung“, AO-StB 9/2018, 291.
- ☞ *Nöcker*, „Quantile in der BFH-Rechtsprechung – Eine erste Analyse der beiden BFH-Beschlüsse X B 16/17 und IV B 4/17 zur sog. Quantilsschätzung“, NWB 40/2017, 3050.
- ☞ *Nöcker*, „Update Betriebsprüfung – Neues vom BFH und die Auswirkungen für die Praxis“, NWB 42/2016, 3157.
- ☞ *Nöcker*, „Zeitreihenvergleich im Gleitschlitten versus Programmierprotokolle der Registrierkasse – Anmerkung zum BFH-Urteil X R 20/13 und einiges mehr“, NWB 48/2015, 3548.
- ☞ *Nowotzin/Teutemacher*, „Die neue Wunderwaffe der Betriebsprüfung: Die Summarische Risikoprüfung (SRP) Das müssen Sie als Steuerberater wissen“, IWW BBP-Sonderheft 2017.
- ☞ *Posch*, „Digitale Ziffernanalyse – Eine Erweiterung des Newcomb-Benford-Verfahrens“, StBp 2010, 338.
- ☞ *Pump*, „Die Kassenführung im bargeldintensiven Betrieb am Beispiel des Friseurs (§§ 158, 162 AO) – Kassenführung mit Barerlösen im bargeldintensiven Betrieb“, StBp 2016, 131, 169, 199, 262, 289, 327 u. 364.

-  *Pump*, „Die Einzelaufzeichnungspflicht von Barerlösen bei offenen Ladenkassen bei Umsatzsteuer und Einkommensteuer – Kryptische Rechtsprechung und der Einsatz von Registrierkassen“, StBp 2015, 1.
-  *Pump*, „Der teure Verzicht auf Verwendung einer Registrierkasse – Beratungsfragen für steuerliche Berater und ihre Mandanten“, DStZ 18/2014, 648.
-  *Pump*, „Die Sicherung des Steuer- und Sozialversicherungsaufkommens bei Taxiunternehmern durch § 30 AO – Integrierte Sicherheitslösung (INSIKA) versus Manipulationssoftware und andere Betrügereien“, VR 11/2013, 361.
-  *Pump*, „Rechtsfolgen bei Verwendung von Manipulationssoftware (Zappersoftware) – Steuerrecht, Steuerstrafrecht und besonderes Strafrecht“, DStZ 9/2013, 299.
-  *Pump/Heidl*, „Hilfestellung bei der Kassenführung durch den steuerlichen Berater – Ärger wegen § 162 AO durch unzulängliche Kassenführung“, StBp 2014, 162 u. 204.
-  *Rau*, „Statistisch-mathematische Methoden der steuerlichen Betriebsprüfung und die Strukturanalyse als ergänzende Alternative“, Reihe: „Steuern, Wirtschaft und Recht“, Bd. 316, Eul-Verlag, 2012.
-  *Schmitz/Gebbers/Huber*, „Nachruf auf Wilhelm Baselt, ‚Vater‘ der modernen steuerlichen Prüfungstechnik“, StBp 2011, 93.
-  *Scholz*, „Der Quantensprung für die Schätzung – Die Quantilsschätzung: Ein Bestandteil der Summarischen Risikoprüfung SRP“ WP Praxis 9/2018, 281.
-  *Scholz*, „Überlegungen des BFH zu Schätzmethoden – Die Quantilsschätzung nach dem BFH-Beschluss vom 12.7.2017“, NWB-BB 11/2017, 351.
-  *Scholz*, „Einsatz der Zeitreihenanalytik in der Prüfungspraxis – heute – BFH-Rechtsprechung im Kontext zur Summarischen Risikoprüfung SRP“, BBK 12/2016, 603.
-  *Schützeberg*, „Die Schätzung im Besteuerungs- und im Strafverfahren“, StBp 2009, 33.
-  *Schumann*, „Urkundenvorlagepflicht und Datenzugriff in Bezug auf IFRS-Abschlüsse und Aufzeichnungen“, Ubg 2018, 400.
-  *Schumann*, „Voraussetzungen und Verfahren der abgaberechtlichen Schätzung nach § 162 Abs. 1 und 2 AO – insbesondere im Fokus der steuerlichen Außenprüfung“, Dissertation im Fachbereich Rechts- und Wirtschaftswissenschaften der Johannes Gutenberg-Universität Mainz, 2009.
-  *Schumann/Wähner*, „Die Quantilsschätzung – Schätzung anhand eines objektivierten Leistungsfähigkeitsmaßes“, Stbg 12/2012, 535.
-  *Schwenkert*, „Erkennen Sie den Unterschied? – Ersetzendes Scannen“, DATEV magazin 2/2015, 16.
-  *Sosna*, „Statistische Ziffernanalyse“, StBp 2004, 249 u. StBp 2005, 97.

-  *Sosna*, „Einsatz statistischer Methoden zur Risikoanalyse, Recherche und Lokalisierung von Steuerausfällen“, StBp 2000, 41 u. 68.
-  *Teutemacher*, „Handbuch zur Kassenführung – Praxishandbuch für die rechtssichere Umsetzung“, NWB-Verlag, 2. Auflage 2018.
-  *Teutemacher*, „Die Kassennachschau kommt! – Die neue Form der ‚Feldüberwachung‘ durch die Finanzverwaltung“, BBK 24/2017, 1160.
-  *Teutemacher*, „Zulässigkeit einer Schätzung mittels Zeitreihenvergleichs – Auswirkungen der jüngsten BFH-Rechtsprechung auf die Betriebsprüfung“, BBK 11/2016, 544.
-  *Teutemacher*, „Aufbewahrung digitaler Unterlagen bei Bargeschäften – Checkliste für eine ordnungsgemäße Kassenführung“, BBK 23/2012, 1073.
-  *Wähnert*, „Das Beweismaß des Besteuerungsverfahrens – Warum eine Konkretisierung des ‚Beanstandungsanlasses‘ aus § 158 AO unverzichtbar ist“, StBp 2018, 199.
-  *Wähnert*, „SRP-Glossar: Wichtige Begriffe zur Summarischen Risikoprüfung (SRP) und zur Neuen Prüfungstechnik (NiPt)“, Stand: Dezember 2017, downloadbar unter:
<https://taxtech.blog/2018/03/09/glossar-zur-summarischen-risikoprufung-srp-und-neuen-prufungstechnik-nipt/>.
-  *Wähnert*, „Schätzungen mit Hilfe von Zeitreihenanalysen – Schrittweise Beschreibung der Quantilsschätzung“, StBp 2017, 323.
-  *Wähnert*, „Zeitgemäße Datenanalyse der Betriebsprüfung – Das Datenprüfungsnetz ‚Summarische Risikoprüfung (SRP)‘“, DB 45/2016, 2627.
-  *Wähnert*, „Engmaschige Prüfungsnetze – Prüfmethode der Finanzverwaltung“, DATEV magazin 4/2016, 15.
-  *Wähnert*, „Die Unterschiedlichkeit von Zeitreihenmodellen – Wie sehr sich Zeitreihenvarianten in der Prüfungsanwendung unterscheiden“, StBp 2016, 61.
-  *Wähnert*, „Logarithmische Normalverteilung als Prüfungsmittel für Wirtschaftsdaten – Die Logik eines leistungsfähigen Verteilungsmusters“, StBp 2016, 1.
-  *Wähnert*, „Eine Antwort auf Petersens ‚wissenschaftshistorische‘ Betrachtung des Benford’s Law – Sachgerechte Maßstäbe für die Neue Prüfungstechnik“, Stbg 12/2015, 511.
-  *Wähnert*, „Verteilungsbasierte Schätzungen im Steuer(straf)recht – Die Überlegenheit von Verteilungen gegenüber Kennzahlen“, StBp 2015, 92.
-  *Wähnert*, „Auch digital ist legal – GoBD“, DATEV magazin 2/2015, 14.
-  *Wähnert*, „‚Bildgebende Verfahren‘ in der (Groß- und Konzern-)Betriebsprüfung – Gläserne Massendaten für korrekte rechtliche Bewertungen“, StBp 2014, 217.

- ☞ *Wähnert*, „Der Beweiswert der Neuen interaktiven Prüfungstechnik (NiPt) – Chi-Quadrat-Test und vergleichbare Verfahren in der Betriebsprüfung“, BBK 8/2014, 380.
- ☞ *Wähnert*, „Deutung von Zeitreihenauffälligkeiten – Die Stabilität wirtschaftlicher Abhängigkeiten“, StBp 2014, 97.
- ☞ *Wähnert*, „Interaktive Außenprüfung – Möglichkeiten der gemeinsamen Sachverhaltsaufklärung im digitalen Zeitalter“, Stbg 1/2014, 20.
- ☞ *Wähnert*, „Aussagekraft und Belastbarkeit von Zeitreihenvergleichen – Erkenntnisse der Neuen interaktiven Prüfungstechnik (NiPt)“, BBK 9/2013, 420.
- ☞ *Wähnert*, „Manipulationspsychologie – Kann Manipulation dauerhaft unsichtbar bleiben?“, StBp 2013, 102.
- ☞ *Wähnert*, „Die Ziffernanalyse als Bestandteil zeitgemäßer Prüfungsansätze“, StBp 2012, 241.
- ☞ *Wähnert*, „Beweiswert neuer Verprobungsmethoden – Gegenüberstellung alter und neuer Ansätze der Betriebsprüfung“, NWB 34/2012, 2774.
- ☞ *Wähnert*, „Das große Potenzial von Strukturvergleichen in der Revision“, StC 4/2012, 18.
- ☞ *Wähnert*, „Betriebsprüfung im digitalen Zeitalter – Untersuchungsgrundsatz versus Mitwirkungsverpflichtung“, BBK 7/2012, 317.
- ☞ *Wähnert*, „Neue interaktive Prüfungstechnik (NiPt)“, Protokoll zum Bochumer Steuerseminar für Praktiker und Doktoranden vom 27. Januar 2012 (Klesen), downloadbar unter: http://www.fachanwalt-fuer-steuerrecht.de/pdf/Protokoll_20100036.pdf.
- ☞ *Wähnert*, „Mehrfachindizien in der Betriebsprüfung – Die unterschätzte Bedeutung von Indizienketten in der Prüfungspraxis“, BBK 24/2011, 1187.
- ☞ *Wähnert*, „Die Beweislast im Besteuerungsverfahren – Brennpunkt ‚Einnahmenvollständigkeit‘“, StC 11/2011, 16.
- ☞ *Wähnert*, „Die Bedeutung von Indizienketten im Besteuerungsverfahren“, StBp 2011, 269.
- ☞ *Wähnert*, „Neue Prüfmethode – Ausfallrisiko mindern“, DATEV magazin 3/2011, 36.
- ☞ *Wähnert*, „Beweislast und Schätzung im Zusammenhang mit modernen Verprobungsnetzen“, StBp 2011, 107.
- ☞ *Wähnert*, „Moderne Einnahmehinterziehung – (ein) Hauptproblem der Besteuerung“, StC 4/2010, 25.
- ☞ *Wähnert*, „Aktuelle Fragen aus der Praxis der Außenprüfung: Digitale Manipulation – Bedrohung für das Steueraufkommen“, StBp 2010, 56.
- ☞ *Wähnert*, „Von zeitgemäßer Risikofilterung bis zu Quantilsschätzung“, StBp 2010, 1.
- ☞ *Wähnert*, „Integration der neuen Prüfungsmethoden in ein systematisches Risikomanagement – Einsatzmöglichkeiten neuer Prüfverfahren für die Finanzverwaltung“, BBK 15/2009, 738.

- ☞ *Wähnert*, „Die Quantilsschätzung – Schätzungsmöglichkeiten mit summarischen Methoden“, BBK 20/2008, 1083.
- ☞ *Wähnert*, „Anwendbarkeit, Aussagekraft und Grenzen stochastischer Manipulationstests“, StBp 2008, 312.
- ☞ *Wähnert*, „Summarische Risikoprüfung – Echtes Risikomanagement in der Finanzverwaltung mit Hilfe einer neuen Prüfungsmethode“, StBp 2008, 296.
- ☞ *Wähnert*, „Die betriebswirtschaftliche und statistische Analyse als Mittel der Außenprüfung“, Stbg 2007, 289.
- ☞ *Wähnert*, „Die Anwendung von Wahrscheinlichkeitstests in der Außenprüfung und die richtige Interpretation der Ergebnisse“, StBp 2007, 65.
- ☞ *Wähnert/Pump*, „Problemfragen der Schätzung nach § 162 AO“, Protokoll zum 140. Bochumer Steuerseminar für Praktiker und Doktoranden vom 4.3.2016 (Wilms), downloadbar unter: http://www.fachanwalt-fuer-steuerrecht.de/pdf/Protokoll_20100118.pdf.
- 📖 *Webel*, „Steuerfahndung – Steuerstrafverteidigung“, NWB-Verlag, 3. Auflage 2016.
- ☞ *Webel/Danielmeyer*, „Schnittstellenverprobung elektronischer Betriebsverwaltungen – das kommende Standardprüffeld?“, StBp 2015, 353.
- ☞ *Webel/Wähnert*, „Summarische Risikoprüfung: Einleitung von Steuerstrafverfahren in der digitalen Welt“, PStR 3/2015, 70.
- ☞ *Webel/Wähnert*, „Der Verdacht einer Steuerstraftat während der Außenprüfung – § 10 BpO im Kontext der Digitalisierung“, NWB 44/2014, 3324.
- ☞ *Weiskopf/Wähnert*, „Neue interaktive Prüfungstechnik – Transfereffekte moderner Verprobungsansätze auf die Großbetriebsprüfung“, BBK 3/2012, 127.
- ☞ *Weiskopf/Wolenski*, „Prüfungsansätze moderner, interaktiver Verprobungsverfahren in der Konzern- und Großbetriebsprüfung“, StBp 2013, 33.
- ☞ *Werner/Wähnert*, „Verbildlichende Prüfung am Beispiel der Gewinnverlagerung – Die Überlegenheit der Visualisierung von Datenregeln und -verteilungen gegenüber Kennzahlen“, IStR 19/2016, 811.
- ☞ *Wiggen*, „Die Beweiskraft des Zeitreihenvergleichs“, StBp 2008, 168.
- ☞ *Wolenski*, „Das Zeitreihenurteil des X. Senats des BFH vom 25. 3. 2015“ u. „Antwort auf Kulosas Replik zur Kritik am Zeitreihenurteil des X. Senats des BFH“, Stbg 6/2016, 268 u. 272.
- ☞ *Wolenski/Wähnert*, „Systematische Visualisierung von internationalen Gewinnverschiebungen – Beispiele für verbesserte Möglichkeiten durch Neue interaktive Prüfungstechnik (NiPt)“, IWB 3/2015, 105.

-  *Wolenski/Weiskopf*, „Praxisfall: Einsatz indirekter Prüfungsmethoden in der Großbetriebsprüfung am Beispiel der Einzelwertberichtigung“, BBK 3/2013, 128.