

Protokoll zum
Bochumer Steuerseminar für Praktiker und Doktoranden
Vom 27. Januar 2012*

Diplom-Finanzwirt Andreas Wähnert

Finanzamt Kiel-Nord

Außenprüfung

„Neue interaktive Prüfungstechnik (NiPt)“

*Dipl.-Juristin, Dipl.-Finanzwirtin Diana Klesen, Wiss. Mit. am Lehrstuhl für Steuerrecht, Bochum.

Einleitung	2
1. Hinterziehungs- und Fehlerrisiken	3
1.1 Digitale Manipulation	3
1.2 Erfahrungen aus der Prüfungspraxis	4
1.3 Das deutsche Steueraufkommen	5
1.4 Beweisführung im Besteuerungsverfahren	5
2. Prüfungsoptionen und -qualität	5
2.1 Arbeitsweise moderner Verprobungsnetze	5
2.2 Betriebswirtschaftliche Verprobung	6
2.2.1 Vorbehalte gegen die Zeitreihenanalyse	7
2.2.2 Fehlerquellen bei der Zeitreihenanalyse	7
2.3 Stochastische Manipulationstests	8
2.4 Gesamtbeurteilung redundanter Verprobungsergebnisse	9
2.5 Prinzip der Indizienketten	10
3. Beweislastverteilung im digitalen Zeitalter	10
4. Diskussion	10
4.1 Bedeutung für § 158 AO	10
4.2 DavoS-Manipulationsoption als Wettbewerbsvorteil?.....	11
4.3 Typologische Betrachtung von Unternehmen - Risikobewertung	13
4.4 Schätzung infolge Verwerfung der Buchführung	14
4.5 Steuerloch und Risikomanagement.....	15

Einleitung

Einleitend stellt *Herr Prof. Seer* den Referenden *Herrn Wähnert* vor, der Betriebsprüfer ist und die betriebswirtschaftlichen und mathematischen Grundlagenkenntnisse für die modernen Betriebsprüfungstechniken im Rahmen seines wirtschaftlich ausgerichteten Studiums erworben hat. Inhalt der heutigen Veranstaltung solle auch die verfahrensrechtliche Einordnung moderner Systemprüfungsansätze sein. *Herr Prof. Seer* berichtet von einem Wortbeitrag von *Frau Mack* auf dem Finanzrichtertag in Köln am vergangenen Montag. *Frau Mack* hätte den modernen Verprobungsmethoden keinerlei indizielle Wirkung zukommen lassen wollen, die Indizwirkung wäre vielmehr Wunschenken der Fiskalvertreter. *Herr Prof. Seer* sei nun gespannt darauf, welche verfahrensrechtlichen Ableitungen aus den modernen Betriebsprüfungstechniken zu gewinnen seien.

Herr Wähnert erläutert, dass er mit seinem Vortrag die Besteuerungsrisiken und „Neue interaktive Prüfungstechniken (NiPt)“ im digitalen Zeitalter sowie die Auswirkungen auf die Beweislast (§ 158 AO) anhand von Modellen (SRP und DavoS) und Fallbeispielen vorstellen möchte.

Das Modell **SRP** (=“**S**ummarische **R**isikoprüfung“) stelle ein Verprobungsnetz aus einzelnen Verprobungsschritten dar. Die Abkürzung **DavoS** stehe für „**D**atenzugriff **auf v**orgelagerte **S**ysteme“. Er wolle nicht alle Details erläutern, sondern den Kontext der Verfahren verdeutlichen und deren Auswirkungen auf das gelebte Verfahrensrecht. Er beginnt seinen Vortrag mit zwei

Zitaten zur Besteuerungssituation. Er zitiert zunächst *den ehemaligen Bundespräsidenten Horst Köhler*, der anlässlich der Verleihung des Max-Weber Preises für Wirtschaftsethik im Jahr 2008 in Berlin sprach: „*Wenn eine Mehrheit es für clever hält, das Finanzamt zu betrügen, dann gehen nicht nur Steuereinnahmen verloren, dann geraten auch der Rechtsstaat und das Vertrauen in die demokratische Ordnung in Gefahr.*“ *Herr Wähnert* wolle das Kontrolldilemma darstellen, um die indizielle Intention besser verständlich zu machen. Mit dem Zitat wolle er verdeutlichen, dass bei Steuervollzugsmängeln auch die demokratische Ordnung betroffen sei. Zudem zitiert er Professor Klaus Tipke, Die Steuerrechtsordnung, Bd. III: „*Ein Steuerpflichtiger kann sich zur Steuerhinterziehung auch entschließen, weil er weiß oder glaubt, dass alle oder viele andere es ebenso tun.*“ Er hinterfragt sodann kritisch, wie schlimm die Besteuerungssituation im digitalen Zeitalter „wirklich“ sei. Dies wolle er anhand seiner praktischen Erfahrungen veranschaulichen.

1. Hinterziehungs- und Fehlerrisiken

1.1 Digitale Manipulation

Anfangs weist *Herr Wähnert* darauf hin, dass die Steuerhinterziehungsproblematik sich nicht mehr auf den reinen Bargeldhandel beschränke, sondern die digitale Manipulation an Bedeutung gewinne. Zur Thematik der digitalen Manipulation gibt er einen Überblick über die verschiedenen Kassensysteme¹, die insbesondere im Einzelhandel bei Bargeldgeschäften zu finden sein:

- Bei der sog. PC-Kasse lasse sich nachträglich und komfortabel mit Hilfe eines Codewortes, das per USB-Stick eingelesen werden könne, der Datenbestand ändern und auch Jahre später sei die spurlose Löschung ganzer Datenpools möglich.
- Beim sog. PC-Kassensystem bestehe die Möglichkeit, Fiskalberichte zu löschen und die Datenbanken zu reorganisieren, sodass eine Datenwiderherstellung unmöglich sei. Hinter dieser Möglichkeit der restlosen Löschung stecke eine klare Intention. Zudem gebe es sog. Springerschlüssel (Jumper) mit externen Stammdaten zum unauffälligen Weglassen von Umsätzen.

Ein System der digitalen Manipulation (z.B. von Kunden- oder Patientendaten, Warenwirtschaftssystem, Apothekenbereich etc.) sei bereits in den sog. Vorsystemen enthalten. *Herr Wähnert* zeigt exemplarisch anhand eines Auszuges aus dem Lizenzvertrag eines verbreiteten Handwerkerprogramms die Möglichkeit auf, die von der Software unterstützte fortlaufende Rechnungsnummern durch manuelle Eingabe zu umgehen, Rechnungen per Mausclick nicht automatisch in der Buchhaltung zu buchen und Daten unprotokolliert zu löschen. Es sei möglich, im Angebotsbereich des Programms die Überschrift frei zu editieren und statt „Angebot“ eine für den Kunden vermeintliche „Rechnung“ zu erstellen. Diese laufe außerhalb des Rech-

¹ *Herr Wähnert* empfiehlt zu dieser Thematik einen 7-teiligen Aufsatz von Erich Huber, insb. Teile 2-5 enthielten ein Panoptikum der Manipulationsmöglichkeiten, vgl. *Huber*, StPB 2009, 153-161, 185-195, 217-223, 253-260, 286-293, 317-324, 342-349.

nungsnummernkreises unter einer Auftragsnummer. Werde zudem die private Kontonummer für Überweisungszwecke angegeben, sei die Manipulation problemlos und unauffällig geglückt. Insb. über eine Rechnungsnummern-Lückenanalyse sei diese nicht auffindbar. Damit sei die Möglichkeit der Manipulation im Vergleich zu früher erleichtert, wo Manipulation nur ohne Mitwisser und bei Bargeschäften üblich gewesen sei. Manipulationsmöglichkeiten seien in den (auch namhaften) Programmen angelegt.

1.2 Erfahrungen aus der Prüfungspraxis

Herr Wähnert berichtet im Folgenden von den bisher gesammelten Erfahrungen aus der Prüfungspraxis in Schleswig-Holstein bei über tausend SRP-Prüfungen. Nun sei die SRP-Prüfung auch von Hamburg, Mecklenburg-Vorpommern und Sachsen-Anhalt übernommen worden. *Herr Wähnert* erläutert die SRP-Erfahrungen und dass SRP den Einsatz moderner Verprobungstechnik in interaktiven Netzen meine. Festzuhalten sei, dass mehr als 70% der Fälle im Einzelhandel redundant auffällig seien. Er gibt zu bedenken, dass jedoch erst bei nicht im Gespräch mit dem Steuerpflichtigen aufklärbaren Auffälligkeiten ein Indiz vorliege. 20% der nach SRP geprüften Fälle (Dauer der SRP ca. 1-3 Std.) seien auffällig, dass sie nach Ansicht der Betriebsprüfung schätzungsreif seien. Mehrsteuern (USt und LSt) würden regelmäßig in fünf- und sechsstelliger Höhe bei Kleinst-, Klein- und Mittelbetrieben anfallen. *Herr Wähnert* führt aus, dass ein Steuerloch-Bericht aus Schweden aus dem Jahr 2008² vorliege, aus dem sich eine Verteilung des Steuerlochs (dort ca. 12 Mrd. € ggü 120 Mrd. € Gesamtsteueraufkommen) mit fast 40 % von Micro-Companies, also Kleinstbetriebe, ergebe. Aufgrund der hohen Steuermoral in Schweden würden dort bessere Möglichkeiten der Auswertung bestehen. Beim deutschen Steueraufkommen von ca. 500 Mrd. € werde der Steuerausfall auf über 50 Mrd. € geschätzt. Bei den regelmäßigen Mehrergebnissen nach SRP-Prüfung in fünf- bis sechsstelliger Höhe sei evident, dass das Problem der digitalen Manipulation nicht kleinzureden sei.

DavoS bedeute die Prüfung von Vorsystemen (eKasse, Warenwirtschaft, etc.) Zu den DavoS-Erfahrungen berichtet er, dass fast alle Systeme Manipulationsoptionen und -funktionen aufweisen würden. Nachteil des DavoS sei, dass dies extrem zeitaufwendig sei. „Unmanipulative“ Systeme seien laut Programmiereraussagen unverkäuflich. *Herr Wähnert* sehe ein Problem darin, dass viele manipulative Systeme existieren, die dann sämtliche Benutzer in den Graubereich ziehen würden. Man müsse sich von dem Gedanken verabschieden, dass es vor Manipulationen sichere Größenklassen von Betrieben oder Branchen gebe. Kritisch hinterfragt *Herr Wähnert*, wie groß der Schaden für unsere Gesellschaft sei.

² Zu finden unter: http://www.skatteverket.se/download/18.5cbdbba811c9a768f0c80005186/Report_2008_1B.pdf.

1.3 Das deutsche Steueraufkommen

Das deutsche Steueraufkommen des Jahres 2010³ aus Einkommen- (44 Mrd. €), Körperschaft- (12 Mrd. €) und Gewerbesteuer (36 Mrd. €) betrage insgesamt 92 Mrd. €; das aus Umsatz- (137 Mrd. €) und Lohnsteuer (128 Mrd. €) betrage insgesamt sogar 265 Mrd. €. Damit werde 74% des Steueraufkommens und damit der größere Teil, aus Gründen einer kostengünstigen Verwaltung von über 8,5 Mio steuerlich registrierten Unternehmen für deren Kunden und Angestellte abgeführt (Treuhandverwaltung). Damit unterliege es dem Risiko der steuerlichen Anmeldung (und der Kontrollwirksamkeit). Insbesondere Umsatz- und Lohnsteuer seien nahezu immer von Steuerausfällen betroffen.

1.4 Beweisführung im Besteuerungsverfahren

Im Folgenden stellt *Herr Wähnert* die Reaktion des Verfahrensrechts auf diese Situation dar, insb. anhand der Beweisführung im Besteuerungsverfahren. Das Beweisrisiko (§ 158 AO) stehe in direkter Abhängigkeit vom gelebten Verhältnis des Untersuchungsgrundsatz vs. Mitwirkungsverpflichtung. Schon in der RAO sei erkannt worden, dass keine Vollprüfung gewünscht und möglich sei, da ansonsten die durch das Treuhandbesteuerungsverfahren eingesparten Kosten wieder durch die Prüfungskosten neutralisiert würden. Daher solle sich das Beweisrisiko an der Balance zwischen Untersuchungsgrundsatz und Mitwirkungspflichten ausrichten. Das Beweisrisiko (§ 158 AO) bestimme die Wirksamkeit der Besteuerungskontrolle (13.210 PrüferInnen bei 8.571.515 Betrieben⁴) und sei regelmäßig entscheidend für den Prüfungserfolg. *Herr Wähnert* wirft die Frage auf, was sich durch die digitale Erfassung, Dokumentation und Prüfung (entscheidungserheblich) verändert habe. Dies seien zum einen die dargestellten Hinterziehungs- und Fehlerrisiken, zum anderen die im Folgenden erläuterten Prüfungsoptionen und -qualitäten, die heute digital möglich seien, aber sich noch in stetiger Entwicklung befänden.

2. Prüfungsoptionen und -qualität

2.1 Arbeitsweise moderner Verprobungsnetze

Der Verprobungsansatz bestehe in der Erstellung eines betriebswirtschaftlichen Profils, das vor der SRP erstellt werde und folgendes berücksichtige:

- Abhängigkeiten prüfen, z.B. Wareneinkauf/Warenverkauf (Art und Intensität)
- Wirtschaftlichkeit wird bei der Prüfung vorausgesetzt (Wissen und Erfahrung)
- Manipulationsoptionen kennen (direkt und indirekt)

Daraus werde ein betriebswirtschaftliches Profil erstellt, bevor SRP durchgeführt werde. Wirtschaftliche Größen oder wirtschaftliche Verteilung, die dem Manipulationsrisiko unterliegen,

³ Als Quelle gibt *Herr Wähnert* das Statistische Bundesamt Deutschland, www.destatis.de, an.

⁴ Als Quelle gibt *Herr Wähnert* die Bp-Statistik 2010 aus dem BMF-Monatsbericht Sept 2011 an.

sein dabei hinsichtlich Entwicklung, absoluter Größe, Struktur einzuordnen. Sodann werde anhand eines systematisch-redundanten Verprobungsnetzes u.a. die Entwicklung über den Zeitverlauf analysiert:

- Leistungs-BWA⁵ (redundante Schlüssigkeitsprüfung)
- Stochastik (systematische Manipulationssuche)
- evtl Prüfungsvertiefung

Zusammengefasst ergebe sich daraus eine betriebspezifische Intensivverprobung (Sensitivität und Beweiswert), die innerhalb weniger Stunden eine seriöse Schnelleinschätzung des Steuerfallrisikos und die Einstellung der Prüfung unter Hinnahme eines kalkulierten Restrisikos, die Festlegung risikoorientierter Prüfungsfelder oder die Einschaltung der Steuerfahndung ermöglichen.

2.2 Betriebswirtschaftliche Verprobung

Möglichst unverdichtete Buchungsdaten würden durch IDEA zur Durchführung automatisierter Prüfungsschritte in eine SRP-Excel-Vorlage eingespeist. Die SRP-Vorlage für Leistungs-BWA ermögliche (durch die automatisierten Prüfungsschritte: Quartals-RGAS⁶, Monatszeitreihen [einfach/doppelt skaliert], RGAS-Trendanalyse, Monatsnachfragefunktion, Wochentagsnachfragefunktion, Bestandsrückrechnung) eine Gesamtauswertung. Der Prüfungsablauf ergebe sich aus der SRP-Vorlage. Er gehe von langen Zeiträumen (Überblick gebend) in Relationsprüfungen hin zu Einzelgrößenentwicklungen und Rück-/Hochrechnungen. Aus dem betriebswirtschaftlichen Profil ergebe sich eine Erwartung an die Konstanz der Aufschlagssätze eines Betriebes. Diese werde mit den tatsächlichen Daten verglichen. Diese Vergleichsbetrachtung könne z.B. zunächst quartals-, dann monatsweise erfolgen und zur besseren Übersichtlichkeit grafisch dargestellt werden, um Auffälligkeiten zu erkennen. Durch Interaktivität zwischen Betriebsprüfer und Steuerpflichtigem könne man Auffälligkeiten ggf. sofort aufklären. Andernfalls könne die Gesamtauswertung ggf. als Schätzungsgrundlage in Frage kommen oder eine Dokumentation für tatsächliche Verständigungen liefern. Im Folgenden erläutert *Herr Wähnert* anhand eines Beispielsfalls die Möglichkeiten der betriebswirtschaftlichen Analyse, insb. Aufschlagssatzreihe, Monatserlösverhältnis-Gegenüberstellung, Wochenerlösverhältnis-Gegenüberstellung, Warenbestandsentwicklung. Dabei komme nicht nur die Gegenüberstellung der Größen Wareneinkauf und Umsatz in Frage. Auch andere betriebswirtschaftliche Größen könne man ins Verhältnis zueinander setzen. Keine der Methoden sei dabei allein perfekt, vielmehr müsse man mehrperspektivisch arbeiten, also mehrere Aspekte im Sinne eines systematischen Verprobungsnetzes zusammen betrachten. Die grafische Darstellung helfe dabei, die Massen an Informationen besser auszuwerten.

⁵ Betriebswirtschaftliche Auswertung

⁶ RGAS bedeute Rohgewinnaufschlagssatz.

2.2.1 Vorbehalte gegen die Zeitreihenanalyse

Bei der Zeitreihenanalyse werde die Entwicklung von Erfolgsgrößen (z.B. Einkauf/Verkauf) beurteilt. Vorbehalte gegen die Zeitreihenanalyse würden insbesondere aus der Literatur hervorgebracht. Diese seien:

- fehlende Berücksichtigung von Bestandseffekten
- Auswirkung von normalen Erfassungsungenauigkeiten

Dem entgegnet *Herr Wähnert*, dass die Chronologie der Darstellung Bestandseffekte wiedergebe; diese könnten zusätzlich manuell eingepflegt werden. Nur massive Zuordnungsfehler würden die verdichteten Größen signifikant beeinflussen. Zudem sei zu beachten, dass bei der Zeitreihenanalyse die Entwicklung zweier Größen im Kontext betrachtet werden solle. Ggf. gleiche sich ein überhöhter Wert im Zeitverlauf aus. Auffälligkeiten, die sich nicht im Zeitverlauf ausgeglichen hätten, würden zudem gemeinsam mit dem Steuerpflichtigen der Aufklärung bedürfen.

2.2.2 Fehlerquellen bei der Zeitreihenanalyse

Als mögliche Fehlerquellen bei der Zeitreihenanalyse nennt *Herr Wähnert* die Periodenwahl, die Größenanpassung und konjunkturell bedingte Preisschwankungen.

- Periodenwahl:

Kriterien für die Periodenwahl sei grundsätzlich die wirtschaftliche Erwartung, betriebliche Modalitäten (z.B. Einkaufs- oder Abrechnungsrhythmus) sowie Erfassungsgenauigkeit. Eine kurze Periode sei zwar grundsätzlich genau, aber auch fehler- und unschärfenanfällig, eine lange Periode hingegen ggf. ungenau. Es sei daher eine mehrperspektivische Erfassung geboten. Anhand des Beispiels einer Zeitreihenanalyse für den Ein- und Verkauf verderblicher Lebensmittel veranschaulicht *Herr Wähnert* das Vorgehen der zunächst wochenweisen Warenzeitreihenanalyse und weiteren Verdichtung zur monatswisen Warenzeitreihenanalyse.

- Größenanpassung:

Probleme könnten darin liegen, dass Vergleichsgrößen so unterschiedlich seien und der Verlauf der kleineren nicht (richtig) zu erkennen sei. Anpassungsmöglichkeiten seien in der Multiplikation (einer oder beider Größen) oder einer prozentualen Darstellung (Regelfall) zu finden.

- Konjunkturbereinigung

Starke konjunkturell bedingte Preisschwankungen seien, so veranschaulicht *Herr Wähnert* anhand der Dieseldkosten bei Taxiunternehmen, abbildbar. Relevant sei das Problem insb. wenn eine Preisbindung die wirtschaftliche Reaktion auf Kostensteigerungen (z.B. Dieseldkosten) verhindere. Als Lösung bietet *Herr Wähnert* die Kostenbereinigung durch

Anpassung mittels Preisindex (z.B. Dieselpreis) an, welcher einer Mengenverprobung entspreche.

Damit schließt *Herr Wähnert* den Überblick über die betriebswirtschaftlichen Analysemöglichkeiten ab.

2.3 Stochastische Manipulationstests

Die SRP-Vorlage Stochastik ermögliche eine digitale Ziffernanalyse (→ Kompletzmanipulation), Schichtung (→ Prüfbarkeit), Schichtungsvertiefung (→ Nachfragefunktionen), Verteilungsanalyse (→ Spitzenkappung) und Beschränkbarkeit auf Teildatenpools (→ gezielte Suche von Manipulationsspuren). *Herr Wähnert* gibt sodann einen Überblick über die Systematik der Verprobungswege, wobei er die betriebswirtschaftliche Verprobung den stochastischen Manipulationstests gegenüberstellt:

- Außerbetrieblicher Vergleich:

Bei der betriebswirtschaftlichen Verprobung werde ein Vergleich von Erfolgsgrößen (z.B. RGAS) mit Erfahrungswerten (Richtsatzsammlung) vorgenommen. Bei den stochastischen Manipulationstests erfolge ein Vergleich der Verteilungsstrukturen in Risikodaten mit Erfahrungswerten (z.B. Benford's Law)

- Innerbetrieblicher Vergleich:

Die betriebswirtschaftliche Verprobung sehe eine Beurteilung der Entwicklung betrieblicher Erfolgsgrößen (z.B. Zeitreihenvergleich) vor. Die stochastischen Manipulationstests würden eine Gegenüberstellung der Verteilung von Teildatenbeständen (z.B. Jahre oder Wochentage) vornehmen.

Für die betriebswirtschaftliche Verprobung stelle sich die Problematik, wie die Bewertung des Unschärfebereichs erfolgen solle. Stochastische Manipulationstests würden die Zufallsstreuung einzelner Ziffern in der Gesamtheit der Buchungen berücksichtigen (χ^2 -Test⁷). Es werde getestet, ob Daten rein zufällig entstanden seien. Sofern dies nach dem Test nicht der Fall sei, müsse geprüft werden, ob dafür eine betriebliche Besonderheit oder Manipulation ursächlich sei. Die Stochastik nimmt den Vergleich von Verteilungen in den Blick. Anhand von Daten eines Eiscafés verdeutlicht *Herr Wähnert* im Folgenden das Vorgehen bei stochastischen Manipulationstests, insb. des „ χ^2 -Tests“. Sofern eine Ziffer in einem homogenen Teildatenpool häufiger auftrete, als dies nach dem Zufallsprinzip zu erwarten sei, müsse diese Auffälligkeit geklärt werden. Allerdings erläutert *Herr Wähnert*, dass der χ^2 -Test umso unempfindlicher werde, je weniger Daten einbezogen würden. Als Beispiel führt er an: Werfe man eine gleichgeprägte Münze zehn Mal, so sei eine Verteilung von 2:8 hinsichtlich der oben liegenden Seiten statistisch möglich. Werfe man die Münze jedoch 10.000 Mal, sei ein Ergebnis von 2.000 zu 8.000 statistisch nicht

⁷ Sog. „Chi-Quadrat-Test“.

zu erwarten. Sofern der „ χ^2 -Test“ Auffälligkeiten eines Teildatenbestandes dahingehend ergebe, dass dieser nicht per Zufall entstanden sein könne, müsse erforscht werden, ob es dafür eine plausible Begründung gebe. Zu beachten sei, dass nach dem „Benford’schen“-Gesetz bzgl. der ersten beiden Zahlenpositionen in Wirtschaftsdaten nicht das Gleichverteilungsgesetz bestehe (z.B. Wahrscheinlichkeit bei Erstziffer: eins=30%, zwei=18%, neun=4,5% -Dies finde sich in der „Log-Nor-Verteilung“ wieder). Davon abweichende Wahrscheinlichkeiten könnten aus betrieblichen Gründen gelten. In der Praxis gebe es aber auch Manipulationen im Sinne einer Wunschbesteuerung dahingehend, dass Steuerpflichtige gesamte Tagesergebnisse aus dem Kassenbericht löschen und durch andere (kleiner) Beträge ersetzen. Die sich dabei manifestierenden Zahlenvorlieben des Manipulierenden lägen außerhalb des Zufalls und könnten mit diesem Test ausfindig gemacht werden. Noch innerhalb des Zufalls könne das manipulierte Ergebnis allerdings liegen, sofern von bspw. 170 Geschäftsvorfällen die letzten 30 gestrichen würden.

Herr Wähnert zeigt im Folgenden ein Prüfungsschema zur Vermeidung von Fehldeutungen auf. Zunächst erfolge eine Ziffernanalyse (digZA). Sofern diese Auffälligkeiten aufweise, liege der Verdacht einer kompletten Manipulation (Schattenbuchführung) nahe. Betragsuntersuchungen seien dann nur bedingt aussagekräftig. Wenn die Ziffernanalyse hingegen unauffällig sei, erfolge eine Betragsschichtung, bei Auffälligkeit sodann eine Verteilungsanalyse. Sofern diese unauffällig sind, liege kein Verdacht auf mit Stochastik begründbare Manipulationen vor. Bei Auffälligkeit hingegen liege ein rechnerischer und grafischer Hinweis auf Verkürzung hoher Einnahmebeträge vor. Anhand eines Diagramms zur Verteilungsanalyse verdeutlicht *Herr Wähnert*, dass Abweichungen von der sog. „Log-Norm-Verteilung⁸“ aus manipulationspsychologischer Sicht bedenklich seien. Sie würden auf ein ungewöhnliches Ereignis hindeuten. Wenn dieses nicht erklärbar sei (z.B. Kapazitätsgrenze erreicht, daher keine Umsatzsteigerung), werde ein Indiz für eine Manipulation daraus.

Problematisch sei die Frage, ob die „Log-Norm-Verteilung“ wissenschaftlich erhärtet werden könne. Die Erfahrungen der Wirtschafts- und Betriebsprüfer und die volkswirtschaftliche Analyse aufgrund von „sicheren“ Daten würden jedoch nach Ansicht *Herrn Wähnerts* eindeutig auf die „Log-Norm-Verteilung“ hinweisen.

2.4 Gesamtbeurteilung redundanter Verprobungsergebnisse

Für eine Gesamtbeurteilung müssten die Ergebnisse der betriebswirtschaftlichen Verprobung und der stochastischen Analyse zusammengeführt werden. Die sich ergebenden ungewöhnlichen Ereignisse müssten für die SRP im Kontext beurteilt werden. Entscheidend sei im Folgenden die substantiierte Mitwirkung des Steuerpflichtigen durch belegbare und quantifizierbare Begründung. Erfolge diese nicht, bestehe eine Schätzungsnotwendigkeit (§ 162 II 2 AO). Die Gesamtbeurteilung veranschaulicht *Herr Wähnert* anhand einer Beispielsauswertung eines SRP-Falls.

⁸ Kurzform für logarithmische Normalverteilung.

Herr Prof. Seer bittet *Herrn Wähnert* die Vorteile der „Log-Norm-Verteilung“ und die daraus zu gewinnende Erkenntnis noch einmal zusammenfassend darzustellen. Die Logarithmische Normalverteilung, so *Herr Wähnert*, sei eine Referenz für die Beträge von verschiedenen Wirtschaftsdaten. Damit könne man die Wirtschaftsdaten auf ihre Zusammensetzung hin überprüfen, nicht auf den absoluten Wert. „Log-Norm“ sei eine einfache Verteilung, die sich nur aus Mittelwert und Standardabweichung zusammensetze. Daraus könne jeder Punkt der Verteilung errechnet werden. Auffälligkeiten würden sofort ins Auge springen. Der Vorteil im Gegensatz zur Ziffernanalyse liege darin, dass es hierbei um echte Beträge gehe. Ergebe die Ziffernanalyse, dass bspw. die „acht“ zu oft vorgekommen sei, helfe dies allein noch kaum weiter bei der Frage nach dem „warum“. Bei der Log-Norm-Verteilungskurve aber könne man sehen, dass bspw. die hohen Umsatzbeträge zugunsten der kleinen beschnitten worden seien.

2.5 Prinzip der Indizienketten⁹

Jede Auffälligkeit habe eine Zufallswahrscheinlichkeit (korrespondierend mit Unschuldswahrscheinlichkeit) und eine Gegenwahrscheinlichkeit (Schuldswahrscheinlichkeit). Dabei erläutert *Herr Wähnert*, dass das Prinzip der beweisverstärkenden Indizienketten laut Rechtsprechung¹⁰ zulässig und auch in anderen Rechtsbereichen üblich sei. Der BGH habe erkannt, dass in ihrer Gesamtwürdigung mehrere Indizien einen Beweiswert hätten, den sie als einzelne Indiztatsachen nicht aufweisen würden. Dass eine Menge an Auffälligkeiten in *einem* Datensatz vorkommen, habe multiplikativen Wert und reduziere die Gesamtunschuldswahrscheinlichkeit erheblich.

3. Beweislastverteilung im digitalen Zeitalter

Voraussetzungen für eine ausgewogene Beweislastverteilung seien:

- Kenntnisse zur elektronischen Ersterfassung, Weiterleitung und Dokumentation (Übertragung der Ordnungsanforderungen, Manipulationsmöglichkeiten)
- Korrekte Bewertung der Ergebnisqualität von Indizien(ketten) aus systematischen Verprobungsnetzen und Vorsystemanalysen
- Kenntnis der umfangreichen Optionen zur gemeinsamen digitalen Aufklärung von Unschlüssigkeiten (Interaktivität).

4. Diskussion

4.1 Bedeutung für § 158 AO

Herr Prof. Seer beginnt die Diskussion mit der Bedeutung der gewonnenen Erkenntnisse für den § 158 AO. Dieser sei Schutzschild der bilanzierenden Unternehmer, der im Sinne eines Vertrau-

⁹ Vgl. dazu *Wähnert*, StBp 2011, 269.

¹⁰ BGH v. 25.11.1982, NStZ 1983, 133.

ensvorschusses die Vermutung der Gesamtrichtigkeit der Buchführung beinhalte. Die Vermutung der materiellen Richtigkeit des Gesamtergebnisses aufgrund der formellen Ordnungsgemäßheit der Buchführung könne jedoch ggf. durch die mehrdimensionale Betriebsprüfungstechnik, die zum Ergebnis der Unplausibilität komme, erschüttert werden. Nach der Rechtsprechung sei zwar der „ χ^2 -Test“ allein nicht ausreichend, um die Vermutung des § 158 AO zu widerlegen. *Herr Prof. Seer* weist jedoch darauf hin, dass entsprechende Rechtsprechung zu mehrdimensionalen Prüfungsmethoden fehle. Der Wortlaut des § 158 AO spreche davon, dass „nach den Umständen des Einzelfalls kein Anlass ist, ihre sachliche Richtigkeit zu beanstanden.“ Nach älterer Rechtsprechung habe dafür das Gegenteil bewiesen werden müssen. Danach wäre, so *Herr Prof. Seer*, eine bloße Gegenwahrscheinlichkeit ehr zu wenig gewesen. Allerdings müsse nach seiner Auffassung im Rahmen des § 158 AO nicht das Gegenteil bewiesen werden; vielmehr reiche der sog. „Gegenbeweis“, also eine Zerrüttung der Vermutung des § 158 AO, aus. Insoweit stimme er *Herrn Wähnert* zu.

Herr Prof. Seer fragt jedoch, was nach der Zerrüttung der Vermutung und des Vertrauensvorschusses des § 158 AO geschehen müsse. Das Finanzamt dürfe seiner Meinung nach gem. § 162 AO unter Auswertung aller betrieblichen Daten schätzen. Die Anforderungen an die Genauigkeit habe sie dabei im Sinne einer „Sphärenorientierten Beweisrisikoverteilung“ abhängig von der Mitwirkung des Steuerpflichtigen zu beachten. Sofern Auffälligkeiten festgestellt worden seien, sei sodann im Sinne einer „Ping-Pong-Theorie“ der Steuerpflichtige am Zug und müsse diese plausibilisieren. Ein Problem ergebe sich dann, wenn der Steuerpflichtige auf die festgestellten Auffälligkeiten nicht reagiere.

Aus dem *Publikum* wird angemerkt, dass § 158 AO die Vermutung der materiellen Richtigkeit des Gesamtergebnisses an die Voraussetzung der formellen Ordnungsgemäßheit der Buchführung nach den Vorschriften §§ 140-148 AO geknüpft sei. Damit müsse die Buchführung auch den Anforderungen des § 146 AO genügen, der eine Unveränderbarkeit der Buchführungsdaten fordere. Sofern man jedoch- wie im Vortrag ausgeführt- von der im VORSYSTEM angelegten Manipulationsmöglichkeit ausgehe, sei die Unveränderbarkeit der Buchführung von Vornherein nicht gewährleistet und die Vermutung des § 158 AO dürfe nicht greifen. Probleme würden darin bestehen, wenn die Finanzbehörde den Nachweis für die Manipulation führen müsse, die Manipulationsmöglichkeit aber auf „zigster“ ProgrammierEbene angelegt sei.

4.2 DavoS-Manipulationsoption als Wettbewerbsvorteil?

Herr Prof. Seer greift die Frage auf, ob bei Manipulationsmöglichkeit bereits keine ordnungsgemäße Buchführung vorliege und § 158 AO somit nicht greife. Er gibt zu bedenken, dass Steuerpflichtige nicht zwingend um die Manipulationsfähigkeit des Programms wüssten. Der Verlust des Vertrauensvorschusses gem. § 158 AO sei jedoch wegen des Wettbewerbsnachteils ggf. ein geeignetes Mittel für die gewünschte Verdrängung derartiger Programme vom Markt. Kritisch

sieht er jedoch die Fälle, wenn keine Auffälligkeiten in den Daten vorliegen, aber wegen Verwendung von Programmen mit Manipulationsmöglichkeit der Vertrauensvorschuss nach § 158 AO verloren gehen sollte.

Aus dem *Publikum* wird angemerkt, dass es keine manipulationsfreien Programme geben könne („Chaos-Computer-Club-Argument“), es sei lediglich eine Frage des Aufwandes, also wie leicht das System Manipulationen zulasse.

Aus dem *Publikum* wird gefragt, wie eine Wiederherstellung des durch Verwendung eines manipulationsfähigen Programms zerstörten Vertrauensvorschusses i.S.d. § 158 AO bei fehlenden Auffälligkeiten möglich sei. *Herr Wähnert* antwortet daraufhin, dass Abweichungen erklärt werden könnten.

Herr Müller merkt an, dass sich in Ansehung des Chaos-Computer-Club-Arguments die Frage stelle, welcher Aufwand genug sei um ein Programm als manipulationsfest i.S.d. §§ 146, 158 AO gelten zu lassen. Er schlägt eine Zwangszertifizierung der Programme (als manipulationsfest) vor, die dann eine grundsätzlich rechtssichere Anwendung des § 158 AO für den Programmnutzer gewährleisten würden. Aus dem *Publikum* wird dazu angemerkt, dass das INSIKA-Verfahren insoweit ausreichend gewesen wäre.

Herr Prof. Seer sieht den Gesetzgeber bei den VORSYSTEMEN mit Hinweis auf die Stornomöglichkeiten im Zugzwang. Er sehe zwei Alternativen: Entweder die vorgeschlagene Zertifizierung vorzunehmen, oder einen Haftungstatbestand (i.S.d. §§ 69 ff. AO) für die Softwarehersteller zu schaffen. Dabei sei es eine Frage der Fahrlässigkeit, ob die Manipulationsmöglichkeit auf tiefergehender ProgrammierEbene bestehe, oder aber auf erster Ebene vom Hersteller selbst als Hinweis gegeben werde. Er halte es jedoch für den elegantesten Weg das Problem de lege lata über die Auslegung der §§ 146, 158 AO zu lösen.

Aus dem *Publikum* wird geäußert, dass die Veränderbarkeit der Buchführung als schlimmster denkbarer Mangel anzusehen sei. Es lasse sich oft kaum feststellen, wie weit die Manipulation gehe. Für den Schätzungsrahmen gebe es aus den vorhandenen Daten dann keinen Ansatz.

Zudem wird aus dem *Publikum* davon berichtet, dass es in der Praxis „Manipulationsverschleierrungs-Tools“ gebe, die Datensätze vor bevorstehender Betriebsprüfung verändern und dann als nicht manipuliert erscheinen lassen. Dazu entgegnet *Herr Prof. Seer*, dass derartige Praktiken an die Öffentlichkeit gebracht werden müssten. Er halte auch die Werbung der Softwarehersteller mit den Manipulationsoptionen für unerträglich. Er habe allerdings auch im Blick, dass eine Vielzahl von Unternehmen nicht manipulieren wolle und sehe es als kritisch an, dass auch diesen der Vertrauensvorschuss des § 158 AO verloren gehen könne. Gegen den Verlust des Vertrauensvorschusses nach § 158 AO wird aus dem *Publikum* eingewendet, dass dies zu weit gehe. Es müsse vielmehr die „Manipulationsbranche“ angeprangert werden. Ggf. solle die Prüfungsdichte sich anhand der verwendeten VORSYSTEME orientieren. Zudem wird aus dem *Publikum* geäußert, dass 95% der Kassensysteme nicht den Anforderungen des § 146 AO genügen wür-

den. Bei strenger Subsumtion des § 158 AO sei der Verlust des Vertrauensvorschlusses in diesen Fällen konsequent.

Auf die Frage, wie die Praxis auf die Verwendung von Programmen mit Manipulationsmöglichkeiten im Hinblick auf die Anwendung des § 158 AO reagiere, wird aus dem Publikum angeführt, dass keine generelle Verwerfung der Buchführung mit Verlust des Vertrauensvorschlusses nach § 158 AO praktiziert werde. Es werde vielmehr im Einzelfall nach Indizien (als nicht erklärbare Auffälligkeiten) gesucht und nur im Einzelfall die Buchführung verworfen.

Herr Prof. Seer meint, dass bei Nichtanwendung des § 158 AO aufgrund der Verwendung von manipulativen Programmen die Buchführung erst verwerfbar sei, wenn Implausibilitäten gefunden würden. Ansonsten seien die betrieblichen Ergebnisse zu verwenden. Es sei auch dann kein freies Hinzuschätzungsrecht der Finanzbehörde gegeben, sondern diese betrieblichen Daten müssten verwendet werden. Dem Finanzamt sei dabei schon ein großer Spielraum eröffnet, die Position des Steuerpflichtigen immer schlechter. § 158 AO sei ein „Damm“, den *Herr Prof. Seer* nicht vorschnell brechen und damit nicht den generellen Verlust des Vertrauensvorschlusses annehmen wolle. Dafür müssten hingegen gute Gründe vorliegen. Ergänzend wird aus dem *Publikum* angemerkt, dass die fehlende Ordnungsgemäßheit der Buchführung auch unter handelsrechtlichen Aspekten, z.B. Überbewertung, relevant sei und dies in die Überlegungen mit einzu beziehen sei.

4.3 Typologische Betrachtung von Unternehmen - Risikobewertung

Herr Prof. Seer fragt, ob die für den Einzelhandel dargestellten Prüfschritte ohne weiteres auf Konzerne übertragen werden können, oder ob die Ausgangssituation dort nicht eine andere sei. Er gibt zu bedenken, dass Betriebsprüfung und große Konzerne mit Compliance-Abteilungen bzgl. der erstrebten Aufdeckung von Manipulationen in Datensystemen in einem Boot mit der Finanzverwaltung sitzen würden. In der Manipulation der Buchführung durch einzelne Mitarbeiter sehe er einzelne Untreuetatbestände verwirklicht, die auch das Unternehmen im Eigeninteresse aufzudecken wünsche („win-win-Situation“). *Herr Wähnert* sieht die Masse an Informationen in Konzernen als problematisch an. Die NiPt seien nötig für die Vorprüfung im Konzern und um Unterschlagungs- und Untreuefälle ausfindig zu machen.

Von *Herrn Prof. Seer* gefragt, wie nach *Herrn Wähnerts* Erfahrungen bei seinen SRP-Schulungen dessen Kollegen die neuen interaktiven Prüfungstechniken (NiPt) aufnehmen würden, antwortet *Herr Wähnert*, dass es zwar unabhängig von deren Alter Zweifler gebe. Nach einiger Übung mit dem Umgang der NiPt würde deren Verständnis jedoch wachsen und damit die Akzeptanz der NiPt.

Auf die Frage von *Herrn Prof. Seer*, wie die Beraterschaft die Thematik beurteile, werden aus dem *Publikum* Stimmen laut, die das Verwerfen der Kassensysteme durch die Finanzbehörden

auf sich zukommen sehen, ähnlich wie es bei der Fahrtenbuchproblematik der Fall gewesen sei. Kritisch wird hinterfragt, ob dies nicht an der Lebenswirklichkeit vorbeigehe.

4.4 Schätzung infolge Verwerfung der Buchführung

Herr Prof. Seer fragt sodann, was im zweiten Schritt die (konstruktive) Konsequenz sei, wenn die Vermutung des § 158 AO (destruktiv) im ersten Schritt zerrüttet worden sei. Er fragt, wie eine dem richtigen Ergebnis möglichst nahe kommende Schätzung i.S.d § 162 AO erfolgen könne. *Herr Wähnert* fordert dann die Mitwirkungspflichten des Steuerpflichtigen ein. Er verweist auf die Leitgedanken der Schätzung. Jedenfalls dürfe im Schätzungsfall der Steuerpflichtige nicht besser stehen, als der kooperative Steuerpflichtige.

*Aus den betriebseigenen Daten des Stpfl. könne eine Spannweite des „Normalen“ herausgelesen werden. Dazu würden sich -besser als arithmetische Methoden- Prozentränge (Quantile) eignen, die in der deskriptiven Statistik zur Einteilung einer Datenmenge in den Standardbereich, schwache und starke Ausreißer verwendet werden. Übertrage man dabei die Verhältnisse der (homogenen) Standardnormalverteilung, definiere sich der Bereich Mittelwert plus/minus mittlere Abweichung (Standardabweichung) mit dem 16%- und 84%-Quantil. Die 68% der dazwischenliegenden Daten umfasse die „Normalfälle“, außerhalb lägen die schwachen und starken Ausreißer. Diese Erkenntnisse würden bei der Quantilschätzung dazu genutzt, um zwischen dem 20%- und dem 80%-Quantil der betriebseigenen Werte zum monatlichen Aufschlagssatz oder Wareneinsatz den Regelgeschäftsbereich (vorsichtig) festzustellen, also den normalen Betriebsverlauf ohne relevante Ausreißer.*¹¹

Der Schätzungsansatz müsse im Standardbereich erfolgen. Sofern Unsicherheiten bestünden, könne die Schätzung sich am oberen Rand, also einem höheren Prozent-Rang orientieren. Innerhalb eines von der Finanzverwaltung errechneten Gürtels dürfe die Leistungsfähigkeit des Steuerpflichtigen dann im oberen Bereich vermutet werden. Um eine gleichmäßige Schätzung in den verschiedenen Bundesländern zu gewährleisten, halte er eine an den Prozenträngen orientierte Systematisierung für erforderlich.

Herr Dr. Krumm fragt nach der wissenschaftlichen Fundiertheit des „Log-Norm-Verfahrens“. Er sieht ein Problem in der Möglichkeit des empirischen Nachweises. Anders, als z.B. in den Sozialwissenschaften für die Empirie gefordert, gebe es nicht den Idealfall, dass z.B. 500 ehrliche Steuerpflichtige und 500 Steuerpflichtige, die betrogen haben, probenhalber anhand des Programms geprüft werden könnten, da man nicht wissen könne, ob die 500 Steuerpflichtigen wirklich ehrlich seien. Er fragt, wie eine wissenschaftliche Verprobung dann aussehen könne. *Herr Wähnert* betont, dass er das Betragsverteilungsgesetz, also die Log-Norm-Verteilung, für einen elementaren Bestandteil der Analyse halte. Bisher sei die Log-Norm-Verteilung allerdings im

¹¹ Die Ausführungen zu den Schätzungsmethoden im vorgehenden, kursiv gedruckten Abschnitt sind zur besseren Verständlichkeit entnommen aus den Erläuterungen Herrn Wähnerts aus: *Wähnert, StBp* 2011, 107, 110.

Gegensatz zum „Benford’schen Gesetz“ tatsächlich nicht empirisch belegt. Nach seinem Wissen über seriöse Untersuchungen sei die Log-Norm-Verteilung jedoch regelmäßig einschlägig. Zudem werde derzeit in einer interdisziplinären Studie versucht, die Log-Norm-Verteilung als Betragsverteilungsgesetz zu verifizieren. Bedeutsam sei, dass die untersuchten Daten hinsichtlich ihrer Risiken oder Richtigkeit einschätzbar seien. *Herr Dr. Krumm* schlägt vor, sich dafür ggf. der DAX-30-Unternehmen zu bedienen, die aufgrund ihrer Compliance-Strukturen ein geringeres Risiko aufweisen würden. Dem stimmt *Herr Wähnert* zu.

4.5 Steuerloch und Risikomanagement

Herr Prof. Seer kritisiert die derzeitig praktizierte, am Mehrergebnis orientierte, statische Einordnung der Betriebsprüfungsbedürftigkeit nach Größenklassen. Aus der von *Herrn Wähnert* zitierten schwedischen Statistik zur Verteilung des Steuerausfallvolumens habe sich ergeben, dass tendenziell das höchste Risikopotential in Klein(st)betrieben liege. Er sehe ein Erfordernis, die Schwerpunktsetzung anzupassen. Auch in den USA werde die Problematik des „Tax Gap“ und des höchsten Risikopotentials in Klein(st)betrieben gesehen und in der Literatur besprochen. Er kritisiert, dass aber auch in den USA nicht auf die Erkenntnis reagiert werde und stattdessen die höchste Prüfungsdichte bei den Großunternehmen vorliege. *Herr Prof. Seer* sehe die verfassungsrechtlich gebotene Gleichmäßigkeit der Besteuerung (Art. 3 I GG, § 85 AO) gefährdet, wenn man bedenke, dass Kleinstbetriebe (im Unterschied zu Großbetrieben) statistisch nur alle 98 Jahr geprüft werden. Darin sehe er ein strukturelles Vollzugsdefizit und widerspreche damit vehement der aus dem *Publikum* vorgebrachten Kritik der Unsachlichkeit der Debatte. Daraufhin wird aus dem *Publikum* Kritik dagegen geäußert, dass Kleinunternehmen unter Generalverdacht gestellt würden. *Herr Wähnert* hält eine bessere Prävention durch Erhöhung der Steuermoral für nötig, etwa am Beispiel Schwedens, wo Steuerhinterzieher stigmatisiert würden. Zudem sei im Rahmen des Risikomanagements zu beachten, dass Betriebsprüfungen bei Großbetrieben oft nicht in der Aufdeckung echter Steuerausfälle resultieren würden, sondern in der Verschiebung der Steuerschuld in andere Prüfungszeiträume.

Herr Prof. Seer resümiert, dass die Rechtsprechung zur Thematik der NiPt abzuwarten sei und lobt die interdisziplinären Forschungsprojekte der Universität Kiel, an denen *Herr Wähnert* beteiligt sei.

Literaturhinweise:

Huber und Seer, StuW 4/2007 S 355

Huber, StBp 2009 S 153

Härtl und Schieder, StBp 2011 S 33

Wähnert, BBK 24/2011 S1187

Wähnert, SteuerConsultant 11_11 S 16

Wähnert, StBp 2011 S 269

Wähnert, DATEV magazin 3_2011 S 36

Wähnert, StBp 2011 S 107

Wähnert, NWB 37/2010 S 2959

Wähnert, SteuerConsultant 4_10 S 25