

Protokoll zum
140. Bochumer Steuerseminar für Praktiker und Doktoranden
Vom 4.3.2016^{1*}

Andreas Wähnert

Dipl. Finanzwirt, Finanzamt Kiel-Nord

Hermann Pump

Richter am Finanzgericht a. D., Münster

„Problemfragen der Schätzung nach § 162 AO“

* Anna Wilms, Wiss. Mit. am Lehrstuhl für Steuerrecht, Bochum.

| | | |
|---|--|----|
| 1 | Einleitung | 2 |
| 2 | Grundlagen der Schätzung | 2 |
| 3 | Ordnungsmäßigkeit im Wandel der Betriebsverwaltung | 4 |
| 4 | Nachweis- und Beweismaßstäbe | 6 |
| 5 | Methoden zur Plausibilitätsprüfung | 6 |
| 6 | Schätzungsmethoden zur Bewertung der Höhe nach | 11 |
| 7 | Resümee | 13 |

1 Einleitung

Einleitend stellt *Herr Prof. Seer* die Referenten *Herrn Wähnert* und *Herrn Pump* vor. *Herr Wähnert* ist Betriebsprüfer und Entwickler der sog. Summarischen Risikoprüfung (SRP) und *Herr Pump* ist Richter am Finanzgericht a. D.

Mit Blick auf das Vortragsthema führt *Prof. Seer* an, dass Kassen und deren ordnungsgemäße Erfassung heutzutage immer schwerer zu überschauen seien. Bei früheren Gesprächen mit Vertretern des BMF und Steuerpolitikern habe er festgestellt, dass die Kasse dort wenig Beachtung finde. In der parlamentarischen Runde der Haarmann Steuerkonferenz 2016 war jedoch das Gegenteil der Fall. Es wurden sogar hinsichtlich der aktuellen Modernisierung des Besteuerungsverfahrens Überlegungen angestellt, dies um einen Teil zur Ordnungsmäßigkeit von Kassensystemen nachzubessern oder zu ergänzen. Dass dieses Thema nun auch in der Steuerpolitik Beachtung erfahre, sei mehr als beachtlich. Des Weiteren merkt *Prof. Seer* an, dass sich mittlerweile auch der X. Senat des BFH mit den neuen Kalkulationsmethoden auseinandersetze. Zum einen vertrete der X. Senat seit einer fundamentalen Entscheidung aus dem Jahr 1989 eine sphärenorientierte Beweisrisikoverteilung und zum anderen befasse sich eine aktuelle Entscheidung mit dem Zeitreihenvergleich (BStBl. II 2015, 743), welche differenziert versuche diese Methode des Zeitreihenvergleichs in die beweisrechtliche Dogmatik der §§ 158, 162 AO einzufügen.

Prof. Seer gibt das Wort nun an *Herrn Wähnert*, der den Vortrag anhand seiner Präsentation leiten und von *Herrn Pump* ergänzt werde.

2 Grundlagen der Schätzung

Herr Wähnert erläutert, dass er in seinem Vortrag neben der Kassenproblematik auch einen Überblick über aktuelle kritische Aspekte der Schätzung, vor allem hinsichtlich der neuen Prüfungstechnik, geben werde. Die Summarische Risikoprüfung (SRP) biete unabhängig vom Prü-

fungsthema einen Einstieg in die Betriebsprüfung, da sie versuche in kürzester Zeit einen Überblick über eine Vielzahl von Daten zu verschaffen. Die Frage des § 158 AO komme neben normalen Betriebsprüfungen zudem auch in Konzernbetriebsprüfungen vor, namentlich bei Verrechnungspreisen. Auch die Schätzungs spiele im Rahmen der SRP natürlich eine Rolle.

Herr Wähnert führt an, dass die Methodik der Betriebsprüfung von hoher Bedeutung sei. Diese müsse natürlich im Rahmen des Verfahrensrechts liegen. Sie müsse jedoch zunächst klare Sachverhaltsverhältnisse schaffen und habe zu diesem Zeitpunkt noch nichts mit dem Verfahrensrecht zu tun. Aus wahrnehmungspsychologischen Gründen könne dies über Visualisierungen der Daten besser erreicht werden, welche somit auch zu einer besseren Prüfbarkeit des § 158 AO führen. *Herr Wähnert* betont, dass er kontinuierlich an der Verbreitung des SRP-Ansatzes arbeite und seine Arbeit hierzu transparent halte.

Im Anschluss führt *Herr Wähnert* ein Zitat von Brinkmann (Schätzungen im Steuerrecht, 2015, S. 232) an, welches besage, dass der deutsche Gesetzgeber sich mit § 162 AO für die Fälle nicht möglicher Ermittlungen steuerrelevanter Sachverhalte für die Schätzung entschieden habe. Dem sei zuzustimmen, da durch die Alternative des Außerachtlassens nicht berechenbarer oder ermittelbarer Besteuerungsgrundlagen keine gerechte Besteuerung erreicht werden könne. *Herr Wähnert* fasst zusammen, dass das Mittel der Schätzung eine notwendige Konsequenz des Verifikationsprinzips sei, da es ein wesentliches Gebot der Gerechtigkeit sei, die Besteuerung auch gleichmäßig durchzusetzen. Laut BFH vom 16.5.2013 (BStBl. II 2014, 225) sei der Gesetzgeber verpflichtet dieses in ein verfassungsgemäßes normatives Umfeld einzubetten, das die Gleichmäßigkeit der Besteuerung insbesondere durch die Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip gewährleiste. Gem. § 162 Abs. 1 AO sei zu schätzen, wenn Besteuerungsgrundlagen nicht ermittelt oder berechnet werden können. Schätzungsvoraussetzung sei, dass die Finanzbehörde alle Umstände zu berücksichtigen habe, die für die Schätzung von Belang seien. Insbesondere seien Schätzungen vorzunehmen, wenn die Buchführung oder die Aufzeichnungen der Besteuerung nicht § 158 AO entsprechen (§ 162 Abs. 2 S. 2 AO). *Herr Wähnert* erläutert unter Bezug auf *Seer* (T/K, § 162 AO, Rz. 6), dass die sphärenorientierte Beweisrisikoverteilung des § 162 AO den Steuerpflichtigen durch eine Beweislastentscheidung nicht benachteiligen dürfe, wenn er seine Mitwirkungspflichten erfüllt. Zusammenfassend stellt *Herr Wähnert* klar, dass der Mitwirkung von Steuerpflichtigen im Besteuerungsverfahren eine entscheidende Bedeutung hinsichtlich des Auslösens bzw. Nichtauslösens der Schätzungspflicht zukomme und zwar in Form des gewährten Prüfungszugangs und der Aufklärungsunterstützung. Anschließend geht *Herr Wähnert* auf die Voraussetzungen ein, die § 158 AO an den Vertrauensvorschuss sachlicher Richtigkeit von Steuererklärungen knüpfe. § 158 AO besage:

„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.“ Es liege also im Nachweis-/Beweisrisiko des § 158 AO ein Zusammenspiel des Untersuchungsgrundsatzes nach § 88 AO und der Mitwirkungspflicht nach § 90/200 AO vor. Es müssen ordnungsgemäße Unterlagen vorgelegt werden, damit keine Entziehung der Besteuerung durch „Beweisverderber“ entstehe. Das äußere Gesamtbild der Unterlagen sei eine ausreichende Grundlage, da es nicht praktikabel und i.d.R. nicht möglich sei, eine so hohe Vielzahl von Geschäftsvorfällen lückenlos zu überprüfen. Der Beweisverderber, so der BFH, dürfe aus seinem Verhalten keinen Vorteil ziehen, sodass die Beweiswürdigung demnach auch belastende Unterstellungen und nachteilige Schlüsse rechtfertige.

3 Ordnungsmäßigkeit im Wandel der Betriebsverwaltung

In diesem Zusammenhang geht *Herr Wähnert* darauf ein, dass die Buchhaltung früher ausschließlich manuell in Haupt- und Nebenbücher eingetragen wurde und der sachverständige Dritte demnach neben seinen Buchhaltungskenntnissen keine höheren technischen Qualifikationen benötigt habe. Heutzutage werden hierzu relationale Datenbanken im Rahmen des sog. E-Managements (elektronische Betriebsverwaltung) genutzt, die eine Vielzahl von Tabellen, Bild-dokumenten und Ähnlichem enthalten können. Hier sei nun eine häufig unübersichtliche Anzahl komplexer Daten gespeichert, deren Ordnungsmäßigkeitsprüfung ein hohes technologisches Verständnis erfordere. Diese Entwicklung sei heutzutage üblich und aus Gründen der Effizienzsteigerung aus der Wirtschaft heraus initiiert worden.

Der ordnungsgemäße Ablauf dieses E-Managements bestehe, so *Herr Wähnert*, aus einer (Erst-) Erfassung als Protokollierung der betrieblichen Vorgänge durch Verwendung von PLU-/Scannerkassen, das Einpflegen von Warenwirtschaftsdaten, das Scannen von Belegen u.a. Diese Formen der Protokollierung dienen der Prozessbeschleunigung sowie einer vollständigen Erfassung. Anschließend müsse die Möglichkeit bestehen diesen Datenbestand zusammenzufassen, zu übertragen und weiterzuleiten. Hierzu müssen interne Schnittstellen existieren oder Systemwechsel vorgenommen werden können. In einem dritten Schritt müsse aus diesem Datenbestand eine Besteuerungsdokumentation gesichert sein (Buchführung bzw. EÜR). Hierzu seien Funktionen der Datenkonsolidierung, Datenorganisation und Datenarchivierung notwendig. *Herr Wähnert* resümiert, dass die Einhaltung der abstrakten Ordnungsvorgaben zu den jeweilig gewählten Verfahren prozessbezogen beurteilt werden müsse.

Die Formalbewertung der Ordnungsmäßigkeit der Buchführung müsse sich an den tatsächlichen Verwaltungsabläufen orientieren. Der BFH habe entsprechend mit seiner Entscheidung vom 16.12.2014 (BStBl. 2015 II S. 519) festgelegt, dass ein Steuerpflichtiger durch die Verwendung eines modernen PC-Kassensystems, welches Kassenvorgänge umfassend einzeln und detailliert aufzeichne und auch eine langfristige Speicherung derselben ermögliche, die Entscheidung zur Aufzeichnung der einzelnen Verkäufe getroffen habe. Zusätzlich entschied der BFH in 2015 (BStBl. 2015 II S. 743), dass eine lückenhafte Dokumentation zur Kassenprogrammierung hinsichtlich der Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis dem Fehlen von Tagesendsummenbons bei Registrierkassen in seinen Auswirkungen gleichstehe. *Herr Wähnert* stellt heraus, dass die höchstrichterliche Formalbewertung immer mehr auf dem prozessorientierten Verständnis der vorliegenden Aufzeichnungs- und Aufbewahrungsverfahren basiere.

Herr Pump merkt hier an, dass die hiermit zusammenhängenden Probleme eine Großzahl von Steuerpflichtigen betreffe, dass z.B. am FG Münster mittlerweile ein Drittel der Verfahren Schätzungen betreffe. So sollen 70-90 % der Kassenaufzeichnungen und damit der Buchführung so schwer fehlerhaft sein, dass § 158 AO als Richtigkeitsvermutung weg falle und die Beweislast umkippe. Dies habe schwerwiegende Konsequenzen für die Steuerpflichtigen, welche z.B. ihre Konzessionen etc. verlieren können. Jedoch seien Berater solcher Mandanten ebenso gefährdet, da Evidenzhaftung vorliegen könne. So greife die Haftung nicht nur, wenn sich über den Beratungsvertrag eine Verpflichtung für die Kassenführung ergebe, sondern auch wenn der Berater Fehlerquellen erkenne und den Mandanten nicht informiere. In dem letzten Fall werde der Berater als Erfüllungsgehilfe zivilrechtlich schadensersatzpflichtig, sodass der Mandant zum Kläger werde. Eine intensive Auseinandersetzung mit der Beleg- und Einzelaufzeichnungspflicht seitens der Berater sei geboten. Hinsichtlich der Einzelaufzeichnung vertritt *Herr Pump* die Auffassung, dass nur tatsächliche Kleinstbeträge (Centbeträge) als unzumutbar einzustufen seien. Ansonsten seien Einzelaufzeichnungen durchweg notwendig. Dies umfasse auch z.B. Einzelaufzeichnungen von Bekleidungsunternehmen und Floristen. Eine gegenteilig lautende Entscheidung des BFH stuft *Herr Pump* als Systemfehler ein. Berater müssen Risiken in diesen Bereichen erkennen und sich notfalls von ihrem Mandanten trennen, wenn die Gefahr krimineller Handlungen bestehe. Er führt als Beispiel ein Urteil des FG Rheinland-Pfalz an, in dem ein Softwarehersteller durch den Verkauf von Kassenmanipulationssoftware an eine Eisdiele in Haftung genommen wurde. Nur bei dieser einen Eisdiele sei ein Mehrsteuerbetrag von EUR 2,4 Mio. entstanden, für die der Softwarehersteller einstehen müsse. *Herr Pump* bemängelt, dass

weitgehend Risiken und Folgen hinsichtlich der problembehafteten Höhe der Schätzung unterschätzt werden und der Fokus zur Vermeidung dieser Probleme auf den Beginn des Prozesses, also auf die Datenerfassung, gelegt werden müsse.

4 Nachweis- und Beweismaßstäbe

Herr Wähnert setzt seinen Vortrag mit der zweiten Voraussetzung des § 158 AO fort. Dies sei, losgelöst von der Ordnungsmäßigkeit, der Beanstandungsanlass, wonach Buchführung und Aufzeichnungen der Besteuerung zugrunde zu legen seien, wenn hinsichtlich des Einzelfalls kein Anlass bestehe, ihre sachliche Richtigkeit zu beanstanden. Demnach sei die gesetzliche Vermutung des § 158 AO hinfällig, wenn es nach Verprobung usw. unwahrscheinlich sei, dass die dargestellten Verhältnisse mit den tatsächlichen Verhältnissen übereinstimmen (AEAO zu § 158). Der Steuerpflichtige habe die Pflicht Bedenken oder Zweifel an der sachlichen Richtigkeit der Buchführung durch Erfüllung seiner Mitwirkungspflicht gem. § 90 AO zu beseitigen. Die Umsetzung des Nachweisverhältnisses in § 158 AO erfordere laut *Herrn Wähnert* eine Abgrenzung von „Strengbeweis“ und „Beanstandungsanlass“. Ersterer sei ein Geständnis, Augenschein etc. mit einer Beanstandungssicherheit von 99,5% (BGH v. 14.12.1989 – 4 StR 419/89). Dieses Maß entspreche der naturwissenschaftlichen Einordnung (Stochastik = Wahrscheinlichkeitstheorie). Als Konsequenz seien die korrekten Besteuerungsgrundlagen anzusetzen und zwar, falls sie nicht ermittelbar seien, mittels Schätzung gem. § 162 Abs. 1 AO. Demgegenüber bezeichne der „Beanstandungsanlass“ Indizien mit einer größeren (Rest-)Unsicherheit im Sinne von Anscheinsbeweisen gegen die sachliche Richtigkeit. Diese müssen nicht die Beweisqualität erfüllen, so *Herr Wähnert*, wie am Beispiel der Nachkalkulation mit all ihren Informationsdefiziten nachvollziehbar ist. Hier existiere keine eindeutige Kausalität. In anderen Wissenschaftsbereichen und z.B. bei Stichproben seien Signifikanzniveaus i.H.v. 95% gebräuchlich. Aus der speziellen Besteuerungsnorm des § 158 AO resultieren hier erhöhte Mitwirkungspflichten des Steuerpflichtigen, wobei die Pflicht zur Schätzung von Amts wegen bei fehlender Aufklärung/Aufklärbarkeit durch § 162 Abs. 2 S. 2 AO bestehe.

5 Methoden zur Plausibilitätsprüfung

Herr Wähnert erläutert, dass ein rein verbalisiertes Bewerten von Plausibilitätsergebnissen zu starken Verzerrungen führen könne, wie z.B. aus einer Abhandlung zu Schriftgutachten für das BKA bekannt ist (nachzulesen unter: http://www.bka.de/nn_193746/SharedDocs/Downloads/DE/Publikationen/Publikationsreihen/01_PolizeiUndForschung/1__26__ProbabilistischeSchlussfolgerungenSchriftgutachten.html). Bei

der schlichten verbalisierten Gegenüberstellung von Wahrscheinlichkeiten von Hypothesen könne es danach zu Problemen in der Wahrnehmung kommen, wenn den Wahrscheinlichkeiten keine geschätzten Zahlen zugeordnet werden. In der Prüfungspraxis können verschiedene neue Prüfungsmethoden zur Anwendung kommen, wie z.B. der Zeitreihenvergleich, der Chi-Quadrat-Test, das Benfordsche Gesetz oder der Vergleich mit der Logarithmischen Normalverteilung. Zu diesen Methoden gebe es z.T. irreführende Bezeichnungen und Beschreibungen, was eine objektive Beurteilung behindert. Nachfolgend stellt *Herr Wähnert* die Gegenüberstellung von Zeitreihenmodellen anhand herkömmlicher Aufschlagsatzzeitreihen und der doppelskalierten Einzelgrößenzeitreihe der SRP vor. Bei dem herkömmlichen Zeitreihenvergleich werde der Aufschlagsatz eines Betriebes chronologisch fortgetragen. *Herr Wähnert* stellt ein Beispiel vor, in dem Wareneinkauf und Umsatz derart auf einen Faktor reduziert werden. Wegen der Reduktion der Information über zwei abhängige Betriebsgrößen auf deren mathematisches Verhältnis, beispielsweise mittels eines Aufschlagsatzes, könne das Zeitreihenergebnis durch eine Vielzahl Einflüsse (z.B. Bestandseffekte, Konjunktur etc.) beeinträchtigt werden. *Herr Wähnert* bevorzuge daher als zentrale SRP-Zeitreihe einen doppelskalierten Einzelgrafenvergleich, welcher eine robuste Wiedergabe der Handlungslogik, vergleichbar einer Mengenverprobung, mit Feststellung von echten Prüffeldern biete. In dem vorgestellten Beispiel wurde die Aufschlagzeitreihe im Hintergrund mit abgebildet, bei der starke Schwankungen zu sehen sind und nicht erkennbar sei, welcher Effekt (Bestandsveränderung, Konjunktur u.s.w.) diese verursache. Bei der doppelskalierten Zeitreihe mit größenangepassten Y-Achsen (z.B. hier für Einkauf und Umsatz) stehe der Faktor, also der Aufschlagssatz, nicht im Vordergrund. Es werde vielmehr untersucht, ob eine Korrelation zwischen den zwei Zeitreihen Einkauf und Umsatz existiere. An dem Beispiel stellt *Herr Wähnert* die Besonderheit klar, dass bei diesem Modell die Alltagseffekte umfänglich wegfallen. So würden Aufschlagsätze deutlich variieren und Bedenken an der sachlichen Richtigkeit hervorrufen, während die doppelskalierte SRP-Zeitreihe eine plausible Handlungslogik abbilde. Bei ernsthafter Betrachtung des Beanstandungsanlasses und Einbeziehung in die Rechtsbeziehung des § 158 AO müsse sich hiermit auseinandergesetzt werden. Eine kalkulatorische Theorie reiche hier nicht aus. Für die Ergebnisbewertung neuer Prüfungstechniken sei laut *Herrn Wähnert* Methodenverständnis unverzichtbar. Die neue Prüfungstechnik funktioniere schematisch wie jede andere Gesamtbildprüfung auch. In seinem Bewertungsschema stellte er als ersten Schritt die Festlegung eines Prüfungsfokus voran. Dieser könne z.B. aus der „Zeitreihenreaktionen zweier abhängiger Betriebsgrößen“ bestehen. Für die Bewertung neuer Prüfungstechnik nach den aufgelisteten Schritten seien die Kalkulation und Geldverkehrsrechnung der Standardmaßstab und nicht der Strengbeweis. Der Maßstab sei das Bekannte, welches allgemein

dazu geeignet sei § 158 AO zu widerlegen. In diesem Zusammenhang führt *Herr Wähnert* aus, dass jede Gesamtbildprüfung eine Idealerwartung beinhalte, welche z.B. bei der Richtsatzsammlung der mittlere Wert (ohne Rahmen), bei einer Kalkulation das eurogenaue Rechenergebnis (ohne Auf- und Abschläge) und bei einer doppelteskalierten SRP-Zeitreihe das perfekte Aufeinanderliegen der Graphen bilde. Die Idealerwartung habe die höchste Einzelwahrscheinlichkeit für einen Betrieb (einer Branche) und werde bei der Ziffernanalyse der Einerstelle in Tageskasseneinnahmen beispielsweise durch die Gleichverteilung dargestellt. Wenn also eine Gleichverteilung daraufhin überprüft werden solle, ob diese § 158 AO widerlegen könne, so stelle sich die Frage, ob die Gleichverteilung z.B. bei Tageskasseneinnahmen fehleranfälliger als ein eurogenaues Kalkulationsergebnis ohne jede Toleranz sei. *Herr Wähnert* verneint dies. Der Maßstab sei entscheidend und werde zurzeit vermehrt falsch gewählt. Zusätzlich würden bei Gesamtbildprüfungen i.d.R. Ergebnisstreuungen berücksichtigt, indem Informationsdefizite und Methodenungenauigkeiten zum Beispiel durch Richtsatzrahmen und Kalkulationstoleranzen einbezogen werden. Dies spiegele die Funktion der Wahrscheinlichkeitstheorie wider, mit deren Hilfe ein Wahrscheinlichkeitsspektrum in der Gesamtbildprüfung aufgebaut werde. Hier komme bei Häufigkeitsauswertungen wie der Ziffernanalyse der Chi-Quadrat-Test zum Tragen, welcher prüfe, ob die beobachteten Häufigkeiten von Ziffern dicht an der Gleichverteilung liegen oder weit von dieser entfernt seien. Der Chi-Quadrat-Test sei jedoch nicht an die Gleichverteilung gebunden, er sei eine Anwendung der Wahrscheinlichkeitstheorie, die auf Normal- bzw. Binomialverteilungen basiere. Der Chi-Quadrat-Test stelle mit der Wahrscheinlichkeitstheorie schlicht die sicherste bekannte Form der Beurteilung von Abweichungen dar. Praktisch werden somit die vorhandenen Werte mit den erwarteten Werten verglichen und die Abweichung bzw. Übereinstimmung festgestellt. Bei der Beurteilung von tatsächlichen Werten werden also die wissenschaftlichen Erkenntnisse zur mengenabhängigen Streuung berücksichtigt. Natürlich müsse die zutreffende Methode gewählt werden. So sei der Chi-Quadrat-Test bei der Überprüfung von einer Einzelziffer nicht zielführend, weil in diesem Fall die Binomialverteilung zu einem genaueren Ergebnis führe. Des Weiteren müsse der Vertrauensbereich definiert bzw. eine Abweisungsgrenze festgelegt werden. In einem letzten Schritt des Bewertungsschemas stellt sich die Frage, ob der Steuerpflichtige durch seine Dokumentation der tatsächlichen Verhältnisse (z.B. Aufzeichnung von Sonderpreisen) zur Vorsorge bzw. Aufklärung beitragen könne. *Herr Wähnert* betont, dass die neue Prüfungstechnik Möglichkeiten zur gemeinsamen Sachaufklärung biete, welche sich vorher nie ergeben haben. So ermögliche die SRP-Prüfmethodik die Möglichkeiten zur substantiellen Mitwirkung durch (Einzel-)Korrekturen, Untersuchungseingrenzungen, Konjunkturbereinigungen und Datenreihen-Umdatierung. Durch SRP-Prüfung auf-

gedeckte Auffälligkeiten können bei konkreten Hinweisen des Steuerpflichtigen umgehend abgesichert bzw. aufgeklärt werden. Ziffernanalysen haben indiziellen Charakter, sodass mathematische Auffälligkeiten durch den Steuerpflichtigen im Rahmen seiner sphärenorientierten Mitwirkungspflicht zu erklären seien. Mögliche Einlassungen des Steuerpflichtigen, wie Korrekturen von Fehlerfassungen oder zeitliche Konkretisierungen, können hierbei problemlos umgesetzt werden. Die den Kalkulationen zugrunde gelegten in Exceldateien hinterlegten Formeln seien vom Anwender nicht änderbar, um eine Standardisierung zu gewährleisten. SRP-Ausdrucke enthalten zudem alle Informationen, um unter Kenntnis der untersuchten Daten die Ergebnisse zu reproduzieren. Die Erkenntnisse zu betrieblichen Sonderfaktoren lassen sich in interaktiven Prüfungsnetzen grundsätzlich problemlos berücksichtigen.

Herr Pump führt hier an, dass die Vermutung des § 158 AO widerlegt sei, wenn die Kassenaufzeichnungen nicht in Ordnung seien. Schwere formelle Fehler in der Kassensführung eines Betriebes stellen in der Gesamtheit zugleich materielle Fehler dar, mit der Konsequenz, dass die Richtigkeitsvermutung nicht greife und Beweislastumkehr bestehe. Er weist zudem darauf hin, dass 3-8% der Barerlöse jährlich hinzugeschätzt werden dürfen und dass aus Beratersicht zwingend für den Einsatz von Registrierkassen gesorgt werden müsse. Formelle Ordnungsmäßigkeit werde nur durch Einzelaufzeichnungen mit Registrierkassen oder Kassensystemen erreicht.

Herr Wähnert setzt seinen Vortrag fort und betont, dass die gemeinsame Sachaufklärung einfacher geworden sei. Die von ihm dargestellten SRP-Vorlagen dokumentieren die Ergebnisse nachvollziehbar, da die Vorlagen ausschließlich Excel-basiert seien. Der Zweck der SRP sei es auffällige Einzelsachverhalte als Prüffelder zu suchen und diese mit Hilfe des Steuerpflichtigen zu klären. Es werde also auch der Schutz des Steuerpflichtigen gefördert, da so konkrete Fälle angesprochen und geklärt werden können. Entgegen der bisherigen Nennung von Einzelmethoden sei SRP jedoch ein Prüfungsnetz, sodass auffällige Ergebnisse aus mehreren Indizien bestehen. Des Weiteren werden einzelne Anscheinsbeweise weitestgehend gegen betriebliche Sondereinflüsse abgesichert. *Herr Wähnert* fügt hinzu, dass auch der BGH mit Urteil vom 27.4.2010 (1 StR 454/09) es als rechtsfehlerhaft ansehe, "wenn sich das Tatgericht bei seiner Beweiswürdigung darauf beschränkt, die einzelnen Belastungsindizien gesondert zu erörtern 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 ausreichen." Zusammenfassend stellt *Herr Wähnert* heraus, dass das Gesamtbild eines typischen

zu bemängelnden SRP-Falls aus formeller Ordnungswidrigkeit (unheilbar verifikationsvereitelnd) bestehe und mehrere Indizien gegen die sachliche Richtigkeit entdeckt werden.

Prof. Seer merkt hinsichtlich des von *Herrn Wähnert* zuvor angeführten strengen Beweises an, dass der BGH mit dem sog. Anastasiafall über den Vollbeweis geurteilt habe, dass der Richter von dem Ergebnis seiner Entscheidung in der Weise persönlich überzeugt sein müsse, dass theoretischem Zweifel (also dem Restrisiko) Schweigen geboten werde. Die von *Herrn Wähnert* angeführten 99,5% des Strengbeweises empfinde *Prof. Seer* als sehr hoch. Der BGH habe in einem Urteil diesen Wert von 99,5% akzeptiert, aber er habe die Grenze nicht auf diesen Wert festgelegt.

Hinsichtlich des Signifikanzniveaus von 95% erläutert *Herr Wähnert* auf Nachfrage, dass dies das bekannteste und gängigste Signifikanzmaß für Qualitätsstandards, Pharmakologie u.Ä. sei. Mittels Stichprobentechnik werde eine gewisse Anzahl von Werten/Daten auf ein Merkmal getestet und mit der Wahrscheinlichkeitsrechnung daraufhin überprüft, ob die ausgewerteten Ergebnisse eine ausreichende Sicherheit (mit 95%) bieten, sodass das betreffende Merkmal auch auf eine weitaus größere Menge zutrefe. Übertragen auf § 158 AO bedeute dies, dass der Beanstandungsanlass ausreichend sicher sein müsse und bei einem Signifikanzniveau unter 95% der Schutz des Steuerpflichtigen greife.

Eine *Stimme* aus dem Publikum (*Stimme 1*) merkt an, dass es zu problematischen Fällen kommen könne, wenn ein Prüfer mittels SRP eine Auffälligkeit entdecke und den Steuerpflichtigen zur Abwehr der Schätzung zur Aufklärung des Sachverhalts aufrufe und beide Seiten keine Erklärung für die relevante Abweichung finden. Der Prüfer möge eventuell unerfahren sein und erkenne mögliche Gründe für die Abweichung nicht und Seitens des Steuerpflichtigen ergebe sich trotz umfassender Prüfung und Offenlegung des konkreten Sachverhalts keine Erklärung.

Herr Wähnert merkt an, dass dies strengbeweisliches Denken darstelle. Ein auffälliges Verteilungsmuster, welches mit den praxisrelevanten Erklärungen (z.B. feste Veranstaltungspreise als Ursache für die Häufung einer Ziffer an der Einerstelle) nicht erklärt werden könne, sei eine genauso unwahrscheinliche Gestaltung wie ein auffälliges Kalkulationsergebnis. Letzteres besage eben auch nicht, dass der Steuerpflichtige zwingend manipuliert haben müsse und biete auch keine absolute Sicherheit.

Prof. Seer führt an, dass das Signifikanzniveau als geeignetes Widerlegungsmittel zur Überwindung der Sperre des § 158 AO seitens der Finanzverwaltung herangezogen werde. Er schätze die neue Entscheidung des BFH, welche die Anforderung an die formelle Ordnungsmäßigkeit hoch halte, um die Steuerpflichtigen zur Ordnungsmäßigkeit ihrer Aufzeichnungen zu zwingen. Werde die Ordnungswidrigkeit nicht nachgewiesen, so werde es der Finanzverwaltung erschwert die Sperre des § 158 AO durch neue Kalkulationsmethoden zu überwinden.

Herr Pump führt hierzu aus, dass er diese Entscheidung für sehr problematisch halte. In der Entscheidung existieren mehrere problematische Ansatzpunkte in der Kassenführung. Der BFH habe erkannt, dass zwei Kassen nicht korrekt geführt waren. Konsequenterweise hätte der BFH auch erkennen müssen, dass die Z-Bons laut Sachverhalt bei identischen Kassen unterschiedliche Inhalte haben. Demnach sei die Kassenführung offensichtlich manipuliert worden. Des Weiteren hätte der BFH erkennen müssen, dass eine neben einer Registrierkasse geführte Thekenkasse unzulässig sei, da die Erfassung ausschließlich über die Registrierkasse zu erfolgen habe, sofern eine solche existiere.

6 Schätzungsmethoden zur Bewertung der Höhe nach

Herr Wähnert setzt seinen Vortrag fort und stellt die Maßgaben für Schätzungen im Besteuerungsverfahren vor. Laut BFH sei das ermessensleitende Ziel, die Besteuerungsgrundlagen durch Wahrscheinlichkeitsüberlegungen so zu bestimmen, dass sie der Wirklichkeit möglichst nahe kommen (BFH v. 25.3.2015, BStBl. 2015 II, 743). Des Weiteren solle sie schlüssig, wirtschaftlich vernünftig und möglich sein (BFH v. 8.12.1984, BStBl. 1986 II, 226) und dürfe keinen Strafcharakter haben (BFH v. 1.10.1992, BStBl. 1993 II, 259). Auf der anderen Seite könne die genaue Bestimmung der Besteuerungsgrundlage durch Schätzung nur zufällig erreicht werden (BFH v. 8.12.1984, BStBl. 1986 II, 226). Zudem sei das Finanzamt grundsätzlich nicht nur berechtigt, sondern verpflichtet, innerhalb des Schätzungsrahmens an die Obergrenze zu gehen (BFH v. 9.3.1967, BStBl. 1967 III, 349). Letztlich habe der Steuerpflichtige die Unsicherheiten einer Schätzung auch vollumfänglich hinsichtlich der Höhe zu vertreten (BFH v. 26.4.1983, BStBl. 1983 II, 618).

Im Anschluss geht *Herr Wähnert* darauf ein, ob die vom BFH präferierten Schätzungsmethoden auch tatsächlich überlegen seien. Hier seien vorrangig die Geldverkehrsrechnung und die Aufschlagkalkulation zu nennen. Erstere stelle schlicht erhaltene und benötigte Geldmittel gegenüber. Nachteilig sei in diesem Zusammenhang die fehlende Aufzeichnungspflicht für private Geldzuflüsse und die nicht vorhandene Nachvollziehbarkeit wesentlicher Abflusswege. Grund-

sätzlich sei der nach ständigen Prüfungserfahrungen wahrscheinlichste Abflussweg von Schwarzeinnahmen „Schwarzlöhne“ regelmäßig nicht reproduzierbar.

Die Aufschlagkalkulation sei eine interne Nachberechnung der operativen Einnahmen mit den Nachteilen, dass wesentliche Ausgangsgrößen nicht ermittelbar und multiplikative Fehler wahrscheinlich seien (Hebeleffekte). Das BFH-Urteil vom 25.3.2015 (BStBl. II 2015, 743) diskutiere eben diese Hebelwirkung. In Rz. 4 und 77 beschreibe dieses Urteil für den verhandelten Fall unterschiedliche Formen einer Zeitreihenvergleichsschätzung. Zunächst werde erläutert, dass die beschriebene Durchschnittsvariante auf statistisch höchst empfindlichen Wochengrößen basieren würde (Rz. 4). Tatsächlich habe der Prüfer jedoch den Ansatz des höchsten 10-Wochen-Aufschlagsatzes angewendet, welcher robustere Größen zu Grunde lege aber einen unrepräsentativen Ausreißerwert wähle (Rz. 77). Die Wahl dieses Maximalwertes sei hochgradig fragwürdig, da kleinste und höchste Werte i.d.R. unrepräsentativ seien. Diese beiden Schätzungsvarianten mittels Zeitreihenvergleich seien statistisch fragwürdig, so *Herr Wähnert*.

Der Zeitreihenvergleich könne für eine Schätzung verwendet werden, wenn § 158 AO widerlegt sei. Grundsätzlich bilde der Zeitreihenvergleich ein umfangreiches Bild zur Entwicklung des Betriebes ab. Bei Zeitreihenschätzungen sei zu beachten, dass monatliche Aufschlagsätze oder andere Kennzahlen repräsentative Werte des Betriebes darstellen, welche statistischen Gesetzen unterliegen und somit die BWA als Analyse- und Planungsinstrument ermöglichen. Innerhalb dieser Werte können jedoch Extremwerte verzerrt sein, wodurch die Randbereiche oft stärker streuen. Daher ziehe die beschreibende Statistik Boxplots zur Feststellung des Normbereichs und somit der unrepräsentativen Ausreißer heran. Die Quantilsschätzung nutze die festen Rang-Wert-Verhältnisse aller Normalverteilungen im Sinne der besten, bekannten Beschreibung des Streuverhaltens. Grundsätzlich werde aus den gewählten Daten ein „Regelgeschäftsbereich“ ohne starke und schwache Ausreißer abgeleitet. Durch die Abbildung des normalen Geschäftsablaufs werde die Repräsentativität gesichert und mit der obersten Grenze die „objektivierten Leistungsfähigkeit“ festgestellt. Zusammenfassend stellt *Herr Wähnert* klar, dass die Quantilsschätzung alle betrieblichen Merkmale/Besonderheiten enthalte und Extremwerte und die Länge des Prüfzeitraumes nicht zu Verzerrungen führen würden. Diese Zeitreihenanwendung der Quantilsschätzung wird im Rahmen der SRP nicht für die Verprobung verwendet und ist dafür auch nicht geeignet.

Neben der Quantilsschätzung könne auch eine Schätzung mit Häufigkeitsmustern herangezogen werden. Hintergrund sei bei dieser Variante die Rekonstruktion der Nachfrage auf Basis eines großen Teils der weniger manipulationsanfälligen Einnahmedaten. Es erfolge also eine interne Ableitung der wahrscheinlichen Einnahmeverhältnisse in geeigneten Fällen.

Prof. Seer knüpft an die vorherige Diskussion an und führt das Problem an, dass möglicherweise nicht umfassend geschulte Prüfer diese statistischen Methoden als Verhandlungsposition nutzen. Ein denkbarer Reflex der Rechtsprechung, welche diese Methoden mutmaßlich ebenso nicht durchblicke und erfasse, könne demnach sein, dass die schwierigere Nachvollziehbarkeit der Prüfungsmethoden nicht ausreiche, um eine bestimmte Schätzung zu begründen. Ebenso würden sich Juristen vielmehr an formalen Kriterien orientieren, wonach der Steuerpflichtige bei Missachtung der formellen Anforderungen Schätzungen anhand der angeführten Methoden hinnehmen müsse. *Prof. Seer* vermute, dass die Rechtsprechung hinsichtlich der Verwerfung des § 158 AO jedoch zurückhaltend sein werde. Er stimme der Aussage Petzers zu, welcher fordere, dass Elektronik nicht die Herrschaft über das Recht bekommen dürfe. Es sei ein Reflex der Juristen, zu befürchten, dass die von *Herrn Wähnert* dargestellten Wahrscheinlichkeitsüberlegungen dazu führen, dass stets bestimmte Werte basierend auf der höchsten Wahrscheinlichkeit gewählt werden. Hierauf wollen Juristen sich i.d.R. nicht festlegen lassen. *Prof. Seer* vertrete die Auffassung, dass die Elektronik natürlich nicht die Herrschaft über das Recht bekommen dürfe, aber das Recht müsse die Elektronik „einfangen“.

Herr Wähnert stellt klar, dass dieser heute viel diskutierte Fall, in dem die formelle Ordnungsmäßigkeit tatsächlich erfüllt werde und Auffälligkeiten dennoch ermittelt werden, ein rein hypothetischer Fall sei, welcher in der Praxis nicht existiere. In der Realität gebe es in auffälligen Prüfungen immer Ordnungsmängel und es sei stets das Gesamtbild zu betrachten. Hinsichtlich der berechtigten Sorge eines Automatismus in diesem Prüfungsbereich führt *Herr Wähnert* an, dass die Alternative zu SRP die Heranziehung von Makrocontainern sei. Im Gegensatz zu SRP seien letztere automatisch geprägt und bieten weniger Möglichkeiten zu prüfen was genau in dem "Container" geschehe. SRP sei das Gegenteil von automatischem Arbeiten, da der Prüfer entlang einer Prüfungssystematik geleitet werde, die wichtigen Entscheidungen dabei aber selber treffe und tief in ein Prüffeld einsteige. Natürlich verlange SRP zunächst eine intensive Einarbeitung in diese Methodik.

7 Resümee

Herr Wähnert stellt zusammenfassend klar, dass zwischen der Anwendung der jeweiligen Methode als Verprobungs- und als Schätzungsmethode unterschieden werden müsse und betont, dass mit SRP in erster Linie Prüffelder gesucht werden.

Im Gegensatz zu den Maßgaben der Schätzung im Besteuerungsverfahren (siehe S. 11) habe die strafrechtliche Schätzung die Tatbestandsmerkmale der Steuerhinterziehung unter Berücksichti-

gung des Grundsatzes in dubio pro reo zu prüfen (BFH v. 14.8.1991, BStBl. 1992 II, 128). Des Weiteren sei eine Schätzung grundsätzlich auch im Steuerstrafverfahren zulässig. Sie müsse jedoch vom Tatrichter überprüft werden und anhand der abweichenden strafrechtlichen Verfahrensgrundsätze überzeugen (BGH v. 4.2.1992 – 5 StR 655/91).

Letztendlich diene die Schätzung der Verwirklichung der Gesetz- und Gleichmäßigkeit der Besteuerung. § 162 Abs. 1 AO verpflichte die Finanzbehörde, die Besteuerungsgrundlagen zu schätzen, um den Folgen des Non-liquets zu entgehen.

Abschließend zählt *Herr Wähnert* die Voraussetzungen für eine sachgerechte und zielorientierte Diskussion über Schätzungen auf:

- Eine Schätzung nach § 162 Abs. 2 AO erfordere keine strengbeweisliche Klarheit der sachlichen Unrichtigkeit.
- Neue Prüfungsmethoden seien ebenso wie die herkömmlichen am Maßstab „AO/Rechtsprechung“ (nicht am naturwissenschaftlichen) zu bewerten (dies gelte auch im Rahmen von externen Spezialisten-Gutachten).
- Gesamtbilder nach § 158 AO seien in ihrer indiziellen Gesamtheit aus formellem und materiellem Eindruck zu beurteilen.
- Vollsätzungen können die Realität nicht wiederherstellen (Kalkulation und Geldflussanalysen müssen realistisch eingeschätzt werden).
- Für eine Schätzung im Besteuerungsverfahren gelte ein vergleichbarer Zweifelsgrundsatz wie „in dubio pro reo“ nicht.

Herr Pump plädiert für die Entwicklung von Kassenleitfäden und die Bestimmung eines Kassenbeauftragten in jeder Firma, aber auch für die gesetzlich eröffnete Möglichkeit, dass jeder bargeldintensive Betrieb freiwillig durch die Teilnahme am INSIKA - Verfahren oder mit anderen technischen Nachweisverfahren in die Lage versetzt werde, die Ordnungsmäßigkeit seiner Kassenführung und ihre Unveränderbarkeit nachweisen zu können. Die daraus folgenden Vorteile seien durch das Hamburger Verfahren mit der freiwilligen Teilnahme der Taxi – Unternehmer am INSIKA – Verfahren für alle Beteiligten deutlich geworden. Alle Beteiligten hätten dies als vorteilhaft erkannt.

Prof. Seer führt aus, dass er für eine Rückführung des § 158 AO auf seinen Grundgedanken sei. Bei Prüfungen müsse zunächst § 158 AO herangezogen werden. In einem zweiten Schritt sei dem Steuerpflichtigen ein Vertrauensvorschuss zu geben und nachfolgend eine Risikoprüfung

durchzuführen. *Prof. Seer* vertritt die Meinung, dass sog. „enhanced relationships“ auch im Bereich des Mittelstands implementiert werden sollten.

Abschließend bedankt sich *Prof. Seer* bei den Referenten für den Vortrag, schließt das 140. Bochumer Steuerseminar und lädt zum nächsten Bochumer Steuerseminar ein. Das 141. Bochumer Steuerseminar findet am 29.4.2016 im Veranstaltungszentrum der Ruhr-Universität zum Thema „Dispute Resolution - Möglichkeiten und Grenzen der Streitbeilegung im internationalen Steuerrecht“ statt. Referent werden *Prof. Dr. Stephan Rasch* (PwC München) und *Dr. Stefan Greil* (Bundesfinanzministerium) sein.