BERLINER STEUERGESPRÄCHE

97. Berliner Steuergespräch

TAGUNGSBERICHT

Zukunft der Betriebsprüfung in einer digitalen Welt

29. September 2025

97. Berliner Steuergespräch

Zukunft der Betriebsprüfung in einer digitalen Welt

Tagungsbericht von Berthold Welling und Dr. Andreas Richter. Beide Autoren sind geschäftsführende Vorstandsmitglieder des Berliner Steuergespräche e.V.

Berthold Welling | Rechtsanwalt, Geschäftsführer des VCI e.V., Berlin und für Recht und Steuern verantwortlich

Dr. Andreas Richter, LL.M. | Rechtsanwalt, Fachanwalt für Steuerrecht, Partner bei POELLATH, Berlin

GESPRÄCHSLEITUNG

Prof. Dr. Thomas Stapperfend | Präsident des Finanzgerichts Berlin-Brandenburg

REFERENTEN

Lars Wargowske | Referatsleiter im Ministerium der Finanzen und für Europa des Landes Brandenburg

Prof. Dr. Klaus-Dieter Drüen | Lehrstuhlinhaber für Deutsches, Europäisches und Internationales Steuerrecht und Öffentliches Recht an der LMU

Dr. Egmont Kulosa | Richter am Bundesfinanzhof (BFH), X. Senat

PODIUM

Prof. Dr. Christian Dorenkamp | Senior Vice President Tax, Deutsche Telekom AG

Bettina Rodenberg | Rechtsanwältin, Steuerberaterin, Leiterin der Steuer- und Zollabteilung der Henkel AG & Co. KGaA

Dirk Rose | Rechtsanwalt, Steuerberater, Wirtschaftsprüfer, Präsident der Steuerberaterkammer Freistaat Sachsen, Vizepräsident der Bundessteuerberaterkammer

Prof. Dr. *Thomas Stapperfend* begrüßte die Teilnehmenden des 97. Berliner Steuergesprächs zum Thema „Zukunft der Betriebsprüfung in einer digitalen Welt“. Zusätzlich zum eigentlichen Thema der Betriebsprüfung werde Dr. *Egmont Kulosa* die aktuelle BFH-Entscheidung des X. Senats aus der Vorwoche zu den Richtsatzschätzungen näher vorstellen Er übergab das Wort an *Lars Wargowske*.

A. Referate

I. Lars Wargowske[1]

Lars Wargowske ging zunächst auf den rechtlichen Rahmen einer steuerlichen Außenprüfung ein und wies darauf hin, dass diese der Ermittlung des Sachverhalts diene. Die Prüfungsmethoden seien nicht gesetzlich vorgeschrieben, sondern müssten an die Eigenart des geprüften Betriebs, die Umstände des Einzelfalls und die Kenntnisse der prüfenden Personen angepasst werden. Dabei gelte stets der Grundsatz der Verhältnismäßigkeit: Zunächst solle die Methode mit dem geringstmöglichen Eingriff eingesetzt werden, bevor weitere Prüfmethoden zum Einsatz kämen. Kooperatives Handeln zwischen Finanzbehörden und Steuerpflichtigen sei dabei besonders wichtig, unterstützt durch Vertrauensvorschüsse.

Er erläuterte, dass im Land Brandenburg eine gemeinsame Strategie zur Vereinfachung und Beschleunigung des Steuervollzugs entwickelt und über die IHK[2] vereinbart worden sei. Diese Strategie gelte grundsätzlich für alle Fälle. Dabei hob er hervor, dass dies einen deutlichen Fortschritt darstelle, der vor allem auf gegenseitiges Entgegenkommen setze. Zugleich betonte er, dass die Umsetzung am Anfang nicht einfach sei, da alle Beteiligten über gefestigte Vorstellungen verfügen würden und sich ein Stück weit aufeinander zubewegen müssten. Man habe sich auch dazu verhalten, wie Fälle für die steuerliche Außenprüfung auszuwählen seien. Dabei werde ein risikobasierter Ansatz verfolgt, der inzwischen auch in der Abgabenordnung verankert sei. Vor Beginn der Prüfung werde in der Regel mitgeteilt, welche Schwerpunkte gesetzt werden, um den Steuerpflichtigen und deren Bevollmächtigten zu verdeutlichen, warum bestimmte Fragen gestellt werden. Dies erleichtere den Prüfungsprozess, da Fragen im Kontext gestellt werden könnten und nicht unbegründet erscheinen würden. Er erklärte, dass das vorgestellte Modell dazu dienen solle, die steuerliche Außenprüfung in einer schnelllebigen, digitalisierten Welt effizienter zu gestalten.

Anschließend ging er auf die Umstände und Bedingungen der Betriebsprüfung ein. Die maßgeblichen Ermittlungen zur steuerlichen Bemessungsgrundlage würden grundsätzlich vom

[1] Referatsbegleitende Präsentation „Präsentation Lars Wargowske“ abrufbar unter https://www.berlinersteuergespraeche.de/archiv/.
[2] Landesarbeitsgemeinschaft der Industrie- und Handelskammern des Landes Brandenburg.

Steuerpflichtigen selbst durchgeführt; mit Ausnahme von Fällen, in denen die Steuer direkt an der Quelle erhoben werde. Das BVerfG habe dies als Ergänzung des Deklarationsprinzips durch das Verifikationsprinzip formuliert. Damit dies funktioniere, seien Buchführungs-, Aufzeichnungs-, Aufbewahrungs- und Mitwirkungspflichten erforderlich, die den Zugang zu relevanten Informationen für die Außenprüfung sicherstellen.

Die steuerliche Außenprüfung diene der Gleichheit beim Vollzug der Gesetze und unterliege einem Untermaßverbot, das vorschreibe, dass eine bestimmte Anzahl von Prüfungen mit definierter Intensität durchgeführt werden müsse. Es bestehe das Leitbild der Vollprüfung, zunächst Kenntnisse über die Grundgesamtheit zu erlangen, um darauf basierend fundierte Entscheidungen treffen zu können. Dies bedeute jedoch nicht, dass jeder Beleg einzeln geprüft werde; vielmehr müsse geprüft werden, wie dieses Leitbild unter Nutzung digitaler Methoden effizient umgesetzt werden könne. Daraus folge, dass im Rahmen der steuerlichen Außenprüfung eine gewisse Daten- und Informationsasymmetrie bestehe. Ohne Mitwirkung der Steuerpflichtigen hätten die Finanzbehörden zunächst keinen Kenntnisstand im Einzelfall. Dies könne sich in Zukunft durch elektronische Rechnungen und entsprechende Meldesysteme ändern, entspreche jedoch nicht dem Status quo. Steuerlich relevante Daten lägen heute nicht mehr nur in klassischen Systemen wie Finanz-, Anlagen- oder Lohnbuchhaltung vor, sondern zunehmend in vorgelagerten Systemen wie Kassensystemen, ERP-Systemen oder Dokumentenmanagementsystemen. Dies führe zu sehr granularen Daten, die nicht notwendigerweise kuratiert seien, wie in traditionellen Buchführungen üblich. Es sei daher entscheidend, festzustellen, wo sich welche Daten befänden und wie damit umzugehen sei. Rechtlich relevant seien zunächst Aufzeichnungs- und Aufbewahrungspflichten, da nur Daten, die aufzeichnungs- und aufbewahrungspflichtig seien, nach § 147 Abs. 6 AO dem Datenzugriff unterliegen. Freiwillig aufbewahrte Daten seien nach der aktuellen Gesetzeslage nicht zugänglich.

Weiter wies er daraufhin, dass – insbesondere durch die steigende Datenmenge und die wachsende Komplexität der Prozesse bei gleichzeitig steigender Erwartung an die Effizienz – die steuerliche Außenprüfung vor Herausforderungen stehe. Dabei sei es entscheidend, nicht alle, sondern nur die relevanten Daten zu erfassen, was häufig komplex und abstrakt sei. Es sei ein iterativer Prozess zwischen Finanzbehörde und Steuerpflichtigen notwendig, um den erforderlichen Datensatz festzulegen. Als Minimum müssten die in der Steuererklärung angegebenen Beträge erfasst werden. Bei sehr granularen Daten könne dies anspruchsvoll sein, besonders wenn die Prüfung zeitlich verzögert erfolge und die ursprüngliche Datenerstellung nachvollzogen werden müsse. Außerdem stelle die Vielgestaltigkeit von Daten und deren Interpretation eine zentrale Herausforderung in der steuerlichen Außenprüfung dar. Strukturelle Probleme entstünden insbesondere durch Schnittstellen- und Formatbrüche, die nicht aus Absicht, sondern aufgrund der dynamischen Natur der Prozesse und ständiger Veränderungen

entstünden. Ein weiteres Thema sei die Speicherung von Daten im Ausland, die zwar zulässig sei, jedoch zu Schwierigkeiten führen könne, wenn Schnittstellen begrenzt sind und nur eine Teilmenge der Daten abgerufen werden könne.

Diese Aspekte verdeutlichten die Notwendigkeit von Standardisierung, um funktionierende Schnittstellen zwischen Systemen zu gewährleisten. Er verwies auf § 147b AO, in dem man versuche, Schnittstellen zu bestimmen sowie die Lohnschnittstelle DSFinV-K als Schnittstellenbeschreibung für die Kassen. Die Digitalisierung stoße auf die rechtlichen Rahmenbedingungen, was in der Praxis zu Diskussionen führen könne. Hinzu komme die Schnelllebigkeit der Zeit, der Wechsel verantwortlicher Personen und der damit verbundene Know-how-Abfluss, wodurch es schwieriger werde, die Prozesse zuverlässig nachzuvollziehen.

Die Digitalisierung könne in den Systemen, insbesondere in der Buchführung und den nachfolgenden Systemen, dazu führen, dass klassische analoge Fehler verschwänden. Diese beträfen bislang einzelne Belege, bei denen etwas nicht korrekt bearbeitet worden sei. Allerdings könnten neue digitale Fehlerquellen entstehen. Solche digitalen Fehler seien zumeist systemisch und könnten unter Umständen erhebliche Auswirkungen haben, da sie nicht nur einmal auftreten, sondern vielfach wiederholt würden. Deshalb müsse sichergestellt werden, dass solche Fehler rechtzeitig erkannt werden.

Bei der Suche nach digitalen Fehlern würden Massendaten, Datenschnittstellen sowie Prozesse und Algorithmen mit steuerlicher Relevanz in den Blick genommen. Das System solle als Ganzes geprüft werden, um weitestgehend sicher feststellen zu können, ob es funktioniere oder nicht, ohne sich mit zahlreichen einzelnen Sachverhalten auseinandersetzen zu müssen. Entscheidend sei das tatsächliche Funktionieren der Systeme und nicht deren vollständige Beschreibung oder detaillierte Ausforschung. Die Systeme würden zwar vermeintlich immer besser, seien jedoch weiterhin fehleranfällig, da eine vollkommen fehlerfreie Programmierung grundsätzlich nicht möglich sei. KI-Systeme könnten zusätzliche Fehlerquellen erzeugen, etwa wenn Daten ungenau verarbeitet oder Prompts nicht präzise formuliert würden. Prüfer könnten dadurch von der KI erzeugte Snapshots erhalten, deren Korrektheit nicht garantiert sei. Künftig bedürfe es einer zielgenaueren und datenbasierten Prüfung. Möglicherweise ließen sich auch Umsatzsteuerdaten und Meldesysteme nutzen, um kontinuierliche Prüfprozesse zu unterstützen.

Es könne für das Verständnis des Prozesses hilfreich sein, dass derzeit Massendaten vorlägen, die ein Nachbild der Systeme erlaubten, um Tabellenverknüpfungen nachzuvollziehen. Dieses Nachrechnen stelle einen Fortschritt gegenüber summarischen Methoden dar, da dadurch die Validität von Stichproben nicht mehr diskutiert werden müsste. Es gehe primär um die Sachverhaltsermittlung und die Bestimmung der Grundgesamtheit, nicht um die

Ableitung eines rechtlichen Ergebnisses. Wichtig sei dabei eine Transparenz: Steuerpflichtige und deren Bevollmächtigte müssten nachvollziehen können, wie die Daten verarbeitet würden und zu welchem Ergebnis sie geführt hätten. Es gehe also um das Thema Datenrücklieferung.

Weiter führte er aus, was passiere, wenn man weder über Wissen noch Daten verfügen würde. Hierbei könnten willkürliche Entscheidungen entstehen. Verfüge man über Wissen ohne Daten, könne dies zwar als ein „Expertensystem“ funktionieren, sei insgesamt jeoch schwierig. Verfüge man nur über Daten ohne Wissen, sei maschinelles Lernen möglich, aber im Ergebnis ebenfalls problematisch. Ziel sei daher ein hybrides Verfahren, das gute Daten mit fachlicher Expertise kombiniere, um willkürliche Entscheidungen zu vermeiden; wenngleich dies idealtypisch sei. Als mögliches Problem führte er die Gefahr der Vermengung von Korrelation und Kausalität aus. Sowohl die Daten, die der Außenprüfung vorgelegt würden, als auch die Verfahren der Prüfung selbst hätten stark mit Korrelationen zu tun. Insbesondere bei begrenztem Einblick in einzelne Fälle könne man bei annähernd parallelen Verläufen zwar Korrelationen erkennen, dies müsse jedoch nicht kausal sein. Selbst eine sehr hohe Korrelation führe nicht automatisch zu einer einfachen Lösung. Es müsse stets hinterfragt werden, ob eine gefundene Korrelation tatsächlich sinnvoll interpretiert werden könne. Er wolle damit zum Nachdenken über die Vorgehensweise in Prüfungen anregen. Eine weitere Methode, die künftig vermutlich viel diskutiert werde, sei Prozessmining. Dabei könne nachvollzogen werden, ob die Abläufe und Prozesse ordnungsgemäß funktionierten. Dadurch zeigten sich in der steuerlichen Außenprüfung, dass oft Prozesse nicht böswillig, sondern zur Erfüllung von Kundenwünschen angepasst würden. Beispielsweise könnten Rechnungen vorab gedruckt oder verschickt werden, während das System automatisch weitere Rechnungen generiere. Solche Vorgehensweisen führten zu komplexen Fragestellungen und erforderten ein genaues Verständnis der Prozessabläufe. Es existierten zudem weitere Analysemöglichkeiten. So könne mit digitalen Handels- und Geschäftsbriefen nachvollzogen werden, wer wann welche Verantwortung trug. Diese Methoden würden einerseits der steuerlichen Außenprüfung vorgelegt, um Handlungen zu dokumentieren, und andererseits von der Prüfung selbst eingesetzt, um Erkenntnisse zu gewinnen. Zusätzlich bestünde die Möglichkeit, Branchenanalysen durchzuführen, um typisches und atypisches Verhalten zu identifizieren. Werden diese Informationen sinnvoll zusammengeführt, könnten sich Rückfragen an den Steuerpflichtigen teilweise erübrigen, da die Vorgehensweisen bereits dem branchenüblichen Standard entsprechen.

Für die Zukunft der steuerlichen Außenprüfung ließe sich festhalten, dass die Datenqualität eine zentrale Stellschraube sei. Ausreißer in den Daten müssten nicht zwingend als Fehler gewertet werden, und auch der Bias, also die subjektive Bewertung, solle stets berücksichtigt werden. Ziel sei es zudem, die Spürbarkeit der Prüfung für Steuerpflichtige und deren Bevollmächtigte zu verringern, nicht allein durch den Einsatz von Tools, sondern durch einen

durchgängigen End-to-End-Datenprozess. Künstliche Intelligenz könnte dabei als Assistenzsystem eingesetzt werden, sei jedoch aufgrund ihrer Fehleranfälligkeit derzeit untauglich. Er schlug außerdem vor, das Besteuerungsverfahren als einen einheitlichen Prozess zu verstehen, der nicht nur die Deklaration und Verifikation umfasst, sondern bereits bei den Unternehmen beginnt und im Zweifel bis zum finanzgerichtlichen Verfahren reicht. Eine solche Sichtweise würde einen erheblichen Fortschritt bedeuten, da sie die Segmentierung des Verfahrens überwindet. Voraussetzung dafür wäre eine entsprechende Standardisierung. Rechtsnormen müssten hier im Zweifel mehr Spielraum lassen. Digitalfreundliche Gesetze könnten zudem eine Standardisierung und Modularisierung steuerlicher Gestaltungsmöglichkeiten ermöglichen und gestreckte Verfahren am Ende überflüssig machen, sodass das Einreichen der Steuererklärung nicht mehr den Startpunkt des Verfahrens bildet, sondern eher ein Zwischenschritt im Gesamtprozess sei. Abschließend betonte er, dass bei einer Veranstaltung des ifst[3] hervorgehoben worden sei, dass das Vertrauen in die Steuerverwaltung, in deren Fairness und angemessenen Steuervollzug eine zentrale Rolle für Steuerehrlichkeit spiele. Gleichzeitig beeinflusse die Durchsetzungsmacht der Verwaltung die Wahrscheinlichkeit der Steuerhinterziehung: Je höher die Entdeckungs- und Bestrafungswahrscheinlichkeit, desto geringer die Tendenz zur Hinterziehung. Dieses Wechselspiel müsse bei der Gestaltung und Umsetzung des Besteuerungsverfahrens berücksichtigt werden.

Prof. Dr. *Thomas Stapperfend* kommentierte, dass die zentrale Aussage des Vortrags die Sachverhaltsermittlung sei. Aus der Praxis habe er jedoch gelegentlich Zweifel, dass alle Prüfer dies in gleicher Weise sehen. Er erinnerte an frühere Fälle von Quantilsschätzungen, bei denen Prüfergebnisse direkt in Schätzungen übernommen wurden, gegen die der X. Senat interveniert habe. Dabei seien auch die Grenzen der Prüfungsmethoden von Bedeutung, da die Verfahren transparent und nachvollziehbar sein müssten. Anschließend leitete er zum Vortrag von Prof. Dr. *Klaus-Dieter Drüen* über.

II. Prof. Dr. Klaus-Dieter Drüen[4]

Prof. Dr. *Klaus-Dieter Drüen* erläuterte, dass ihm seine erste Erfahrung als Betriebsprüfer vor rund 30 Jahren die Grundlage gebe, über die rechtlichen Aspekte der digitalen Betriebsprüfung zu sprechen. Er wolle sich auf die Rechtsfragen der digitalisierten Betriebsprüfung konzentrieren und mit einem Rückblick auf die Entwicklung beginnen: Von der Reichsabgabenordnung mit reiner Buchprüfung über die AO 1977, die sich dem Einsatz von

[3] Institut Finanzen und Steuern e.V.
[4] Referatsbegleitende Präsentation „Präsentation Prof. Dr. Klaus-Dieter Drüen“ abrufbar unter https://www.berlinersteuergespraeche.de/archiv/.

Datenverarbeitungssystemen öffnete, bis hin zur Einführung der Digitalzugriffsrechte der Finanzverwaltung im Rahmen der Betriebsprüfung im Jahr 2002.

Er fuhr mit der Ausgestaltung der digitalisierten Betriebsprüfung fort. Der normative Ansatzpunkt liege in der Aufbewahrungspflicht des § 147 AO. Die Regelungen seien nicht in den Vorschriften zur Außenprüfung (§§ 193 ff. AO) verankert, was systematisch sinnvoll sei. Die Besteuerung beruhe auf Mitwirkungspflichten, die Aufbewahrungs- und Aufzeichnungspflichten vorsehen würden. Die §§ 140–148 AO regelten die Aufbewahrung als Voraussetzung für die Steuererklärung. Die Digitalisierung sei für Steuerpflichtige freiwillig; sie könnten ihre Aufzeichnungen digital führen, müssten dies aber nicht. Wenn sie dies täten, dürften seit 2002 die Finanzbehörden die Digitalsysteme gemäß § 147 Abs. 6 AO nutzen. Diese Nutzung unterstütze die Gleichbehandlung bei der Steuerverifikation. Trotz verfassungsrechtlicher Kritik, der er sich selbst nicht anschließe, habe sich die Regelung bewährt. Über 20 Jahre Erfahrung zeigten, dass die digitale Betriebsprüfung in der Praxis weitgehend geräuschlos und weniger eingriffsintensiv sei als Vor-Ort-Prüfungen. Änderungen durch den Gesetzgeber hätten lediglich die Datenmitnahme auf Prüferlaptops betroffen.

Das Jahressteuergesetz 2024 stelle klar, dass der digitale Datenzugriff ein Sonderrecht der Außenprüfung sei und für andere behördliche Prüfungen nicht gelte. Bei der Frage, auf welche Daten die Finanzverwaltung zugreifen dürfe, hätten die Gerichte einige Grenzen geklärt: Die gesamte Buchführung, einschließlich Drohverlustrückstellungen, unterliege dem Zugriff, unabhängig von steuerlicher Relevanz. Auch Gewinnermittlungen nach § 4 Abs. 3 EStG seien einbezogen. Außerdem seien E-Mails als Handels- und Geschäftsbriefe vorzulegen, wie am 30.4.2025 vom BFH geklärt.[5] Allerdings dürfe die Finanzbehörde vom Steuerpflichtigen nicht die Erstellung eines Gesamtjournals erzwingen; sie könne nur auf vorhandene Daten zugreifen. Der BFH sehe dafür keine Rechtsgrundlage.[6] Prof. Dr. *Klaus-Dieter Drüen* stimmte dem zu. Gleichstellung durch das Datenzugriffsrecht bedeute nur, dass die Finanzverwaltung auf die vorhandenen Daten Zugriff nehmen kann. Sie sei ein digitaler Trittbrettfahrer. Wenn der Steuerpflichtige gewisse Daten nicht habe, könne er nicht gezwungen werden, diese zu erzeugen. Insoweit habe die Rechtsprechung den in der Literatur entwickelten Gedanken der Prozeduralisierung bestätigt, also die Frage, wie zugriffspflichtige Daten überhaupt zu bestimmen seien. Im Steuerrecht gebe es anders als im Handelsrecht die weit gefasste Vorschrift des § 147 Abs. 1 Nr. 5 AO, wonach alles aufzubewahren sei, was steuerlich von Bedeutung sein könne. Nach der Rechtsprechung könnten hierzu beispielsweise auch Speisekarten von Restaurants zählen, wenn sich daraus Rückschlüsse auf den Wareneinsatz und dessen Plausibilität ziehen ließen. Der BFH habe konsequent entschieden, dass nur das vorlagepflichtig sei, was zugleich aufbewahrungspflichtig sei; die Kontrolle setze somit an der

[5] Beschluss des Bundesfinanzhofs vom 30.04.2025 – XI R 15/23, DStR 2025, 2194.
[6] Ebd. Rn. 43.

Aufbewahrungspflicht an. Freiwillig geführte Unterlagen müssten dagegen nicht vorgelegt werden. Zudem sei die Prozeduralisierung dahingehend fortentwickelt worden, dass der Steuerpflichtige insbesondere bei E-Mails die Erstqualifikation selbst vornehmen müsse. E-Mails könnten privaten oder betrieblichen Charakter haben, aufbewahrungspflichtig oder nicht aufbewahrungspflichtig sein. Der Steuerpflichtige habe daher zu beurteilen, ob bestimmte E-Mails oder andere Daten dem Datenzugriff unterliegen. Die Literatur habe eingewandt, eine solche Möglichkeit der Selektion könne die Verfassungswidrigkeit der Regelung nicht verhindern, wenn der Steuerpflichtige dadurch den Zugriff beschränke. Dem widersprach Prof. *Dr. Klaus-Dieter Drüen*: Wenn der Steuerpflichtige durch eigene Maßnahmen bestimmte Daten vom Zugriff ausschließen könne, liege es in seiner Verantwortung, den Datenbestand nach den maßgeblichen Kriterien zu ordnen. Die Finanzverwaltung müsse anschließend begründen, warum eine solche Selektion fehlerhaft sei und weitere Daten in den Zugriff einbezogen werden müssten. Der BFH habe diese Auffassung in einer jüngeren Entscheidung bestätigt und dabei auch klargestellt, dass die GoBD[7] trotz ihrer praktischen Bedeutung lediglich Verwaltungsanweisungen seien. Im konkreten Fall sei die Vorlage eines Gesamtjournals auf die GoBD und die Grundsätze ordnungsmäßiger Buchführung gestützt worden. Diese Grundlagen seien jedoch keine taugliche Rechtsgrundlage, da Verwaltungsanweisungen keine Rechtspflichten für Steuerpflichtige begründen könnten. Auch dies habe der Bundesfinanzhof in seiner Entscheidung ausdrücklich festgestellt.

Prof. Dr. *Klaus-Dieter Drüen* leitete anschließend zum nächsten Themenkomplex über. Während die Abgrenzung zwischen vorlagepflichtigen und nicht vorlagepflichtigen Daten durch verschiedene Entscheidungen weitgehend geklärt sei, sei die Auswertung digitalisierter Betriebsprüfungsdaten rechtlich noch offen. Nachdem feststehe, welche Daten dem Zugriff der Finanzverwaltung unterliegen, stelle sich die Frage, wie diese ausgewertet werden dürfen und welche Folgen daraus zu ziehen seien. Der Datenzugriff diene als Verifikationsinstrument im Rahmen der Besteuerung. Er sei in die allgemeinen Grundsätze der §§ 85, 88 AO sowie die Prüfungsgrundsätze des § 199 Abs. 1 AO einzubeziehen und bilde als besonderes Mittel der Sachverhaltsermittlung den rechtlichen Schlüssel für digitale Analysemethoden, die ausschließlich im Rahmen der Betriebsprüfung zulässig seien. Zu dieser zweiten Stufe gebe es bislang nur wenige Entscheidungen. Maßgeblich sei die BFH-Entscheidung des X. Senats zum Zeitreihenvergleich, die gezeigt habe, dass digitale Prüfungsmethoden zu Korrekturen steuerlicher Ergebnisse führen können.[8] Dieses Prinzip lasse sich auch auf andere digitale Verfahren übertragen. Seit dieser Entscheidung sei es jedoch ruhig geworden. Verfahren etwa zum Chi-Quadrat-Test seien bislang nicht zum BFH gelangt. Er vermute, dies liege sowohl an der

[7] Grundsätze zur ordnungsmäßigen Führung und Aufbewahrung von Büchern, Aufzeichnungen und Unterlagen in elektronischer Form sowie zum Datenzugriff (GoBD), BMF v. 28.11.2019, IV A 4 – S 0316/19/10003 :001, BStBl. I 1269 (vorherige Version BMF v. 14.11.2014, IV A 4 – S 0316/13/10003, BStBl. I 1450).
[8] Urteil des Bundesfinanzhofs vom 25.03.2015 – X R 20/13, DStR 2015, 1739.

begrenzten Übertragbarkeit der bestehenden Rechtsprechung als auch an einer Zurückhaltung der Finanzverwaltung, um mögliche nachteilige Urteile zu vermeiden.

Dann verwies er auf die aktuelle BFH-Entscheidung, in der es um die Schätzung von Einnahmen auf Grundlage der amtlichen Richtsatzsammlung bei einer offenen Ladenkasse in einer Diskothek ging.[9] Für die Frage nach den Rechtsvorgaben und -grenzen der digitalen Betriebsprüfung lasse sich daraus ableiten, dass der BFH klare Anforderungen an die Qualität der Vergleichsdaten stelle. Eine bloße Bezugnahme auf Fremddaten reiche demnach nicht aus. Diese Maßstäbe müssten auch für digitalisierte Prüfungsverfahren gelten. Da bislang nur wenige gerichtliche Entscheidungen vorlägen und das Gesetz kein Sonderrecht für digitale Betriebsprüfungen kenne, müsse ein solches „Datenzugriffsfolgerecht“ auf Grundlage der allgemeinen steuerlichen Prinzipien entwickelt werden. Ausgangspunkt sei das im Gesetz angelegte System von Deklaration und Verifikation. Die Steuererklärung sei nicht nur Pflicht, sondern auch Recht des Steuerpflichtigen, der als Erster Zugriff auf die Lebenswirklichkeit habe und nach § 150 Abs. 1 Satz 3 AO die Steuer selbst zu berechnen habe. Die Aufzeichnungen des Steuerpflichtigen genössen zudem nach § 158 AO einen Vertrauensvorschuss. Entsprächen Buchführung und Aufzeichnungen den §§ 140 bis 148 AO, seien sie der Besteuerung zugrunde zu legen. Das Gesetz sehe keinen Vorbehalt vor, wonach abweichende Erkenntnisse der Finanzverwaltung Vorrang hätten. Daher genüge auch der Einsatz digitaler Prüfmethoden allein nicht, um von ordnungsgemäßen Aufzeichnungen des Steuerpflichtigen abzuweichen; selbst dann nicht, wenn die digitalen Analysen zu scheinbar „besseren“ Ergebnissen führten. Es ginge nicht um die Frage, welches System es besser könne. Eine Abkehr von der ordnungsgemäßen Buchführung setze stets einen konkreten Anlass voraus. Die Richtigkeitsvermutung nach § 158 AO könne nur erschüttert werden, wenn nachweisbare Fehler oder Unvollständigkeiten vorlägen. Digitale Verfahren allein reichten dafür nicht aus; sie dürften lediglich Indizien, aber keine Beweise liefern. Statistische Methoden wie der Chi-Quadrat-Test oder die Summarische Risikoprüfung könnten Unstimmigkeiten anzeigen, ersetzten aber keine formelle oder materielle Prüfung. Gelinge die Erschütterung der Richtigkeitsvermutung, müsse geprüft werden, welches Verfahren an die Stelle der Buchführung treten könne. Der BFH habe im Urteil zum Zeitreihenvergleich betont, dass eine Methode nur anwendbar sei, wenn sie zum konkreten Betrieb passe, etwa nicht bei atypischen oder stark schwankenden Geschäftsverläufen.

Anschließend widmete er sich den verschiedenen Anforderungen an den Abgleich von Eigendaten mit Fremddaten, die die Finanzverwaltung insbesondere aus Datensammlungen nach § 88a AO gewinnt. Dabei stelle sich die Frage, welche Daten unter welchen Bedingungen erhoben, verarbeitet und verknüpft werden und welche Daten unberücksichtigt bleiben. Letzteres

[9] Urteil des Bundesfinanzhofs vom 18.06.2025 – X R 19/21, DStR 2025, 2300.

könne, wie das jüngste BFH-Verfahren zur Richtsatzschätzung zeige, erhebliche Bedeutung haben. Zentral bleibe die Aussagekraft der Vergleichsdaten für den konkreten Fall. Der BFH habe in seiner Entscheidung zur Richtsatzsammlung klargestellt, dass die Finanzverwaltung zwar nicht-öffentliche Datenbanken zur Ermittlung von Vergleichsdaten nutzen dürfe, diese aber Mindestanforderungen an Datenqualität und Transparenz erfüllen müssten. Damit schließe sich der Senat dem „Armani-Urteil“[10] an und führe aus, dass Gerichte gehalten seien, Nachfragen zur Zusammenstellung und Ableitung der Vergleichsdaten zu stellen. Beim Einsatz von Vergleichsdaten gelte eine besondere Beweislastverteilung. Könnten Fragen zur Qualität der Datensätze etwa aus Gründen des Steuergeheimnisses nicht beantwortet werden, gehe dies zulasten des Beweiswerts und somit zulasten der Vergleichsdaten. Diese Sichtweise sei zutreffend und unstreitig. Die Verantwortung sollte jedoch nicht erst auf der Ebene des gerichtlichen Rechtsschutzes ansetzen. Transparenz und Nachvollziehbarkeit müssten vielmehr bereits im Betriebsprüfungsverfahren gewährleistet sein. Da hier regelmäßig vom Aufzeichnungswerk des Steuerpflichtigen abgewichen werde, bestehe eine Bringschuld der Finanzverwaltung, ihre Datenauswertungen begründet und nachvollziehbar zu erläutern. Anhaltspunkte dafür finde man bereits in der BFH-Entscheidung zum Zeitreihenvergleich. Dort habe das Gericht festgestellt, dass sich die Finanzbehörde durch mathematisch-statistische Methoden in eine „technisch-rechnerische Überlegenheit“ gegenüber dem Steuerpflichtigen versetze.[11] Daraus folge das Gebot verfahrensrechtlicher Waffengleichheit: Digitale Methoden dürften nicht einseitig zugunsten der Verwaltung wirken. Vielmehr seien Transparenz und Überprüfbarkeit auch aus Sicht des Steuerpflichtigen sicherzustellen. Die Datengewinnung, - struktur und - auswertung müssten daher offengelegt werden, um eine sachgerechte Verteidigung zu ermöglichen. Aus § 121 Abs. 1 AO in Verbindung mit der asymmetrischen Datenlage ergebe sich ein besonderes Begründungserfordernis, wenn die Finanzverwaltung vom Buchführungswerk des Steuerpflichtigen abweicht. Es gehe dabei nicht um Fragen des Ermessens wie bei der Prüfungsauswahl, sondern um die Gesetzmäßigkeit der Besteuerung und das System von Deklaration und Verifikation. Der Steuerpflichtige dürfe nicht auf die vermeintliche Überlegenheit digitaler Systeme verwiesen werden. Er zitierte dann Dr. *Egmont Kulosa,* der bereits beim 62. Berliner Steuergespräch 2017 betont habe, dass rechnerische oder grafische Ergebnisse allein noch keine zwingenden Besteuerungsgrundlagen seien, sondern stets in den tatsächlichen Lebenssachverhalt eingeordnet, auf Plausibilität geprüft und rechtlich gewürdigt werden müssten.[12] Dies gelte, so Prof. Dr. *Klaus-Dieter Drüen*, gleichermaßen für digitalbasierte Auswertungen in der Betriebsprüfung, die anstelle der Aufzeichnungen des Steuerpflichtigen herangezogen werden. Digital gewonnene Erkenntnisse seien annahme-

[10] Urteil des Bundesfinanzhofs vom 17.10.2001 – I R 103/00, 171, DStR 2001, 2149.
[11] Fn. 24, Rn. 52.
[12] Vgl. 62. Berliner Steuergespräch vom 6. März 2017, "Herausforderungen der digitalen Betriebsprüfung, unter anderem bei bargeldintensiven Betrieben", Tagungsbericht, S. 2 f., abrufbar unter https://www.berlinersteuergespraeche.de/archiv/.

und wahrscheinlichkeitsgestützt und somit keine zwingenden Argumente. Unterschiedliche Datengrundlagen oder Wahrscheinlichkeitsannahmen könnten zu abweichenden Ergebnissen führen. Ebenso wenig wie eine Besteuerung nach Verdacht zulässig sei, dürfe eine Besteuerung auf Basis eines „Computerverdachts“ erfolgen, alles müsse nachvollziehbar sein.

Die Einwände gegen diese Anforderungen bedeuteten nicht, dass digitalisierte Prüfmethoden unzulässig seien, sondern dass ihre begrenzte Leistungsfähigkeit berücksichtigt werden müsse. Transparenz und Überprüfbarkeit seien das Gegengewicht und zentral für die Wahrung der Waffengleichheit im Prüfungsverfahren. § 32a Abs. 2 AO, der Informations- und Auskunftspflichten einschränkt, betreffe zukünftige Prüfmaßnahmen und nicht die laufende Datenauswertung. Gleiches gelte für Risikomanagementsysteme und § 21a Abs. 1 S. 4 FVG. Daraus folge kein Ausschluss der Überprüfbarkeit; der Steuerpflichtige habe daher ein besonderes Recht auf Einsicht in die digitalen Verarbeitungsvorgänge der Betriebsprüfung.

Er schloss mit Ausführungen dazu, dass der Rechtsrahmen des § 147 Abs. 5-7 AO i.V.m. § 200 Abs. 1 S. 2 AO es der Finanzverwaltung ermögliche, im Rahmen digitalisierter Betriebsprüfungen ermessensgerechte und verhältnismäßige Datenabgleiche vorzunehmen, um den Verifikationsauftrag zu erfüllen. Die Rechtsprechung bestätige weitreichende Zugriffsrechte, habe jedoch punktuell Rechtsgrenzen aufgezeigt. Ein explizites Datenfolgenrecht existiere nicht, die Nutzung und Verwertung der Daten sei nach allgemeinen Regeln, insbesondere § 158 AO, zu beurteilen. Der Zeitraum der maschinellen Aufbewahrung müsse verhältnismäßig sein. Lange Aufbewahrungspflichten im Steuerrecht würden durch § 171 Abs. 4 AO bei Außenprüfungen noch verlängert. Digitalisierung biete einen Impuls für zeitnahe Betriebsprüfungen und spreche für kürzere Aufbewahrungsfristen. Schon 1977 wollte man die Fristen verkürzen, scheiterte aber am Widerstand unter anderem der Betriebsprüfer, die vor der Verjährung warnten. Bis heute sei es nicht zu einer Verkürzung der Fristen gekommen. Dies sei ein Punkt, über den man auch diskutieren müsse.

Prof. Dr. *Thomas Stapperfend* merkte an, dass mit diesem Vortrag bereits ein umfassender Bogen von der Nachvollziehbarkeit der Richtsätze, wie sie die neue Entscheidung des BFH zeige, bis hin zur Übertragbarkeit auf digitale Prüfergebnisse gespannt worden sei. Deshalb sei es nun nicht nur aus Aktualitätsgründen sinnvoll, den Blick von internen Daten zur Außensicht zu wechseln, also darauf, welche Rolle Drittdaten bei Betriebsprüfungen spielten. Diese würden insbesondere dann herangezogen, wenn die eigenen Daten unvollständig seien oder merkwürdige Ergebnisse lieferten. In diesem Zusammenhang kämen häufig die Richtsatzdaten der Finanzverwaltung zur Anwendung. Er wies darauf hin, dass die aktuelle Entscheidung des BFH, die Prof. Dr. *Klaus-Dieter Drüen* erläutert habe, auch für die Diskussion zu digitalen Daten wichtige Aussagen enthalte. Dazu werde im Folgenden Dr. *Egmont Kulosa* Stellung nehmen.

III. Dr. Egmont Kulosa[13]

Dr. *Egmont Kulosa* erklärte, sein Beitrag solle einen ergänzenden Diskussionsbeitrag zu den Ausführungen von Prof. Dr. *Klaus-Dieter Drüen* darstellen. Bereits 2022 habe der X. Senat das BMF um Teilnahme am Revisionsverfahren gebeten und Fragen gestellt; die abschließende Entscheidung sei jedoch erst 2025 nach einer mehrstündigen mündlichen Verhandlung veröffentlicht worden, an der das BMF vorbehaltlos auskunftsfreudig beteiligt gewesen sei. Er wies darauf hin, dass die Ausführungen nicht entscheidungstragend seien, da die Entscheidung der Vorinstanz bereits aus anderen Gründen beanstandet worden sei. Die Ausführungen zur Zulässigkeit der Richtsatzschätzung stellten ein obiter dictum dar.[14] Er stellte klar, dass der X. Senat als einzelner Senat des BFH nicht befugt gewesen sei, die Methodik der Richtsatzsammlung eigenständig zu beanstanden oder zu verändern. Bisher seien alle Senate davon ausgegangen, dass es sich um eine anerkannte Schätzungsmethode handle, deren Grundlagen nie hinterfragt worden seien. Das Hinterfragen dieser Grundlagen stelle daher eine Abweichung von der bisherigen Rechtsprechung dar, sodass der Senat keine tragende Entscheidung treffen konnte. Er erläuterte, dass der X. Senat den Abschnitt zur Richtsatzschätzung in drei Teile gegliedert habe: Erstens die Bestätigung der Richtsatzschätzung als notwendiges Hilfsmittel der Betriebsprüfung, zweitens die kritische Würdigung der heutigen Ausprägung der Richtsatzsammlung, drittens die Begründungsanforderungen für die konkrete Anwendung der Richtsätze.

Der Senat habe betont, dass die Methode nicht verworfen werde und keine Grundsatzkritik geübt werde. Jedoch bestünden Zweifel an der Methodik: Die Richtsatzsammlung sei nicht repräsentativ und genüge damit nicht den Anforderungen der statistischen Wissenschaft.[15] Für Repräsentativität seien eine Zufallsauswahl und eine ausreichende Stichprobengröße erforderlich. Während Letztere bundesweit wohl gegeben sei, fehle es an der Zufallsauswahl. Die Betriebe würden nach den Maßstäben des § 193 AO aufgrund ihrer Prüfungswürdigkeit ausgewählt, ein geeignetes Kriterium für Betriebsprüfungen, aber nicht für eine zufällige Stichprobe. Daher fehle es an Repräsentativität und Methodengerechtigkeit, sodass die Richtsatzsammlung nach Auffassung des X. Senats keine taugliche Schätzungsmethode im statistischen Sinne darstelle.

Außerdem bestünden Bedenken zum Ausschluss bestimmter Gruppen von Betrieben aus der Richtsatzsammlung.[16] Zwar sei das Ausscheiden der jeweils niedrigsten und höchsten 10 %, von Groß- und Kleinstbetrieben sowie von Betrieben mit außergewöhnlichen Verhältnissen

[13] Referatsbegleitende Präsentation „Präsentation Dr. Egmont Kulosa“ abrufbar unter https://www.berlinersteuergespraeche.de/archiv/.
[14] Vgl. Urteil des Bundesfinanzhofs vom 18.06.2025 – X R 19/21, DStR 2025, 2300, Rn. 83 ff.
[15] Ebd. Rn. 89 ff.
[16] Ebd. Rn. 114 ff.

unbedenklich sei; problematisch sei jedoch, dass klare Kriterien fehlten und die Auswahl damit von einzelnen Finanzbeamten abhänge. Hier seien einheitliche Vorgaben durch die Oberbehörden wünschenswert. Zudem könnten ausgeschlossene Betriebe auch nicht anhand der Richtsätze geschätzt werden, da sie nicht in die Datengrundlage eingeflossen seien. Kritisch sehe der Senat insbesondere den Ausschluss von Verlustbetrieben, da dieser die Richtsatzspanne statistisch nach oben verschiebe. Auch die Nichtberücksichtigung späterer Korrekturen im Rechtsmittelverfahren führe zu Verzerrungen, weshalb die Finanzverwaltung einen pauschalen Abschlag von etwa 1–3 % prüfen solle.

Zu den Begründungsanforderungen führte er aus, dass ein zweistufiges Vorgehen erforderlich sei, das aus der Verengung der breiten Richtsatzspanne auf einen engeren Bereich unter Berücksichtigung der betrieblichen Besonderheiten und anschließender Festlegung eines Einzelwerts bestehe.[17] Dabei müssten regionale Unterschiede, Kundenstruktur und Standortfaktoren berücksichtigt werden, insbesondere in der Gastronomie, da Verkaufspreise in Hochpreisregionen deutlich höher seien als in strukturschwachen Gebieten.

Der Senat habe abschließend kritisiert, dass es sich bei der Richtsatzsammlung allgemein um eine ungenaue Schätzungsmethode handle, weswegen sie nachrangig zu innerbetrieblichen Schätzungsanknüpfungen sei.[18]

Er betonte abschließend, dass die Entscheidung zunächst keine praktischen Folgen habe, da es sich um ein obiter dictum handele. Ziel sei eine fachliche Diskussion innerhalb des BFH. Auch der Steuerstrafrechtssenat des BGH sei in diese Diskussion mit einzubeziehen, da dieser auch die Richtsatzschätzung anwendete. Für die Finanzverwaltung bestünden zwei Optionen. Man könne abwarten, insbesondere im Hinblick auf andere Senate des BFH, was jedoch die Gefahr einer jahrelangen Hängepartie berge. Oder man könne auf eine statistisch repräsentative Richtsatzsammlung hinwirken. Die Zufallsauswahl wäre organisatorisch machbar, da die Strukturen für das Zusammenführen der Daten vorhanden seien. Hier müsse nur die Fallauswahl etwas anders organisiert sein als bislang geschehen.

B. Podiumsdiskussion

Prof. Dr. *Thomas Stapperfend* erklärte, dass die Entscheidung an dieser Stelle zwar nicht vertieft diskutiert werden solle, da der Schwerpunkt der Veranstaltung auf der digitalen Betriebsprüfung liege, er aber dennoch von *Dirk Rose* wissen möchte, ob es bereits Überlegungen zu den praktischen Auswirkungen der Entscheidung gebe.

[17] Ebd. Rn. 128.
[18] Ebd. Rn. 136 ff.

Dirk Rose erklärte, die Entscheidung habe unmittelbare praktische Auswirkungen. Er zeigte sich überrascht über die Aussage von Dr. *Egmont Kulosa*, der betont habe, die Richtsatzsammlung solle nicht grundsätzlich infrage gestellt werden. Er verstehe nun, dass dies der innergerichtlichen Diskussion innerhalb des BFH diene, halte die Richtsatzsammlung derzeit jedoch für nicht anwendbar. Er führte aus, dass gerade die „normalen" Unternehmen, auf die in der Entscheidung häufig Bezug genommen werde, in der Praxis meist gar nicht geprüft würden, weil sie ordnungsgemäß arbeiteten und keine Auffälligkeiten zeigten. Daher sei fraglich, wie die Richtsatzsammlung unter diesen Umständen überhaupt als Schätzungsgrundlage dienen könne. Mit Blick auf die zunehmende Digitalisierung regte er an, über eine breitere Datengrundlage nachzudenken, etwa durch eine Ergänzung der Datenbasis über Veranlagungsdaten, sofern dies mit dem Verifikationsprinzip vereinbar sei. Er betonte abschließend, dass der innerbetriebliche Vergleich nach der Entscheidung des BFH Vorrang habe und die Finanzverwaltung künftig stärker auf die tatsächlichen Betriebsdaten zurückgreifen sollte, anstatt sich auf statistische Richtwerte zu stützen.

Prof. Dr. *Thomas Stapperfend* erklärte, der innerbetriebliche Vergleich sei zwar vorrangig, doch gebe es Fälle, in denen ein Rückgriff auf äußere Vergleichsdaten erforderlich bleibe. Aus Beratersicht sollten entsprechende Fälle derzeit offengehalten werden, da noch keine abschließende BFH-Entscheidung vorliege. Er merkte an, die jüngste Entscheidung sei möglicherweise nicht endgültig. Anschließend leitete er zur digitalen Betriebsprüfung in größeren Unternehmen über. Er bat *Bettina Rodenberg* um eine Einschätzung, wie die digitale Betriebsprüfung in der Praxis funktioniere, ob es dabei Probleme gebe und wie nachvollziehbar das Vorgehen der Finanzverwaltung sei.

Bettina Rodenberg erklärte, sie wolle das Thema mit der grundsätzlichen Erörterung ihrer Haltung beginnen. Betriebsprüfung sei für sie vor allem eine Frage des gegenseitigen Vertrauens zwischen Steuerverwaltung und Steuerpflichtigen. Steuerehrlichkeit werde nicht nur von der Verwaltung, sondern maßgeblich auch von den Unternehmen selbst getragen. Sie betonte, bei der *Henkel AG & Co. KGaA* verfolge man einen partnerschaftlichen Ansatz („do ut des") und pflege ein gutes Verhältnis zur Betriebsprüfung. Im Zuge der Digitalisierung bedeute dies, Prüfer aktiv zu unterstützen: Bei Anfragen stehe stets eine Ansprechperson aus der Steuerabteilung bereit, die die Datenprozesse erkläre und so ein besseres Verständnis auf beiden Seiten ermögliche. Sie wies anschließend darauf hin, dass mehr Daten allein keinen Mehrwert darstellten. Gerade bei komplexen Themen wie der globalen Mindestbesteuerung zeige sich, dass Digitalisierung bestehende Systembrüche nicht vollständig ausgleichen könne. Viele Systeme seien gar nicht darauf ausgelegt, die geforderten Daten zu liefern, sodass vollständige Compliance derzeit kaum erreichbar sei. Zudem hob sie hervor, dass die Digitalisierung ein Umdenken verlange, Weg vom Papier und hin zum Prozess. Für Prüfer sei es oft schwierig, betriebliche Abläufe nachzuvollziehen. Zudem müsse die Betriebsprüfung in einer digitalen

Welt deutlich zeitnäher erfolgen, da alte Systeme regelmäßig abgeschaltet oder archiviert würden. Abschließend betonte sie, künstliche Intelligenz werde künftig keine bloße Assistenz mehr sein, sondern als „kommunizierende Röhre“ mit der Unternehmensseite interagieren. In 5 Jahren könnten sich Prüfungs- und Unternehmenssysteme zunächst automatisiert austauschen, bevor der Mensch eingreife. Die Diskussion über die digitale Betriebsprüfung sei daher bereits überholt, die Zukunft habe längst begonnen.

Prof. Dr. *Thomas Stapperfend* schlug vor, die Diskussion über die Zukunft der digitalen Betriebsprüfung zunächst noch etwas zurückzustellen und wandte sich an Prof. Dr. *Christian Dorenkamp* als Vertreter eines Großunternehmens mit der Frage, ob er die geschilderten Probleme ähnlich sehen würde.

Prof. Dr. *Christian Dorenkamp* warf die Frage auf, ob klassische Betriebsprüfungen in der digitalisierten Welt überhaupt noch notwendig seien. Durch die Digitalisierung würden die Unternehmensprozesse und die Buchführung immer besser und fehlerärmer. Klassische Fehlerprüfungen, wie die Korrektur von Buchungsfehlern, Aktivierungs- oder Abschreibungsfehlern, würden dadurch zunehmend überflüssig. Er betonte, dass Fehler durch die Digitalisierung deutlich reduziert würden, wodurch der Aufwand für die Betriebsprüfung in diesen Bereichen kleiner werde. Die Buchhaltungsprozesse, die früher enormen Arbeitsaufwand erforderten, liefen heute weitgehend automatisiert ab, sodass Quartals- oder Jahresabschlüsse inzwischen fast „Non-Events“ für die Accounting-Teams seien. Dennoch sehe er weiterhin einen wesentlichen Aufgabenbereich der Betriebsprüfung, der nicht durch Digitalisierung ersetzt werde, nämlich die Prüfung von Bewertungsunterschieden und rechtlichen Auslegungsschwierigkeiten, bei denen die Finanzverwaltung und die Unternehmen unterschiedliche Auffassungen haben. Dies betreffe insbesondere komplexe Transaktionen, Umwandlungen oder Rechtsfragen, bei denen ein Verfahren bis zum BFH führen könnte. Diese „Grenzfälle“ blieben nach seiner Einschätzung die zentralen Aufgaben der Betriebsprüfung, während die klassischen Fehlerprüfungen durch digitale Prozesse stark reduziert würden. Ein weiterer Punkt betreffe die technische Abwicklung der Betriebsprüfung. Durch digitale Systeme könnten Folgebescheide zwischen Festsetzungsfinanzamt und Betriebsprüfungsfinanzamt heute deutlich effizienter erstellt werden. Dies spare Zeit und vereinfache die Verwaltung erheblich. Er hob hervor, dass zeitnahe Prüfungen für Finanzverwaltung und Unternehmen Vorteile brächten: Die Sachverhalte seien noch aktuell und leichter nachvollziehbar. Alte Systeme, die inzwischen abgeschaltet wurden, stellten kein Hindernis dar. Anschließend empfahl er, die „digitale Dividende“ der Betriebsprüfung zu nutzen. Durch die Effizienzgewinne der Digitalisierung können Altjahre schneller aufbereitet werden, wodurch die Finanzverwaltung bestehende Rückstände verringern und gleichzeitig die Steuereinnahmen steigern könne. Diese zeitnahe Prüfung ermögliche eine gleichmäßigere Besteuerung nach wirtschaftlicher Leistungsfähigkeit, da Sachverhaltsungenauigkeiten reduziert würden. Zudem könnten die zusätzlichen Einnahmen genutzt

werden, um steuerliche Rahmenbedingungen wie die degressive AfA oder Verlustverrechnung gezielt zu gestalten und dadurch fiskalische Spielräume zu schaffen. Sein abschließendes Fazit sei, dass die Digitalisierung die klassische Betriebsprüfung in Routinebereichen überflüssig mache, gleichzeitig aber Kapazitäten für komplexe und rechtlich schwierige Prüfungen schaffe. Sie ermögliche effizientere Abläufe, schnellere Datenzugriffe und eine bessere Nutzung der digitalen Dividende für fiskalische Zwecke. Voraussetzung dafür sei, dass die Prüfungen zeitnah erfolgten und die verfügbaren Ressourcen der Finanzverwaltung effizient eingesetzt würden.

Prof. Dr. *Thomas Stapperfend* wandte sich an *Dirk Rose*, der in erster Linie Mittelstands- und kleinere Unternehmen betreue und nicht die Großunternehmen und fragte, wie sich die digitale Betriebsprüfung aus seiner Sicht gestalte und ob er die Betriebsprüfung ebenfalls künftig als obsolet betrachte.

Dirk Rose wies darauf hin, dass er mehrere Themen ansprechen wolle. Zunächst gebe es Fälle, die „pathologisch“ seien, die Mehrheit der Prüffälle jedoch problemlos und kooperativ abliefen. Er äußerte dann Skepsis gegenüber der von *Bettina Rodenberg* erwähnten KI-basierten „kommunizierenden Röhre“, stattdessen sollte die natürliche Intelligenz des Steuerberaters, der mit dem Betriebsprüfer zusammenarbeitet, weiterhin entscheidend sein. Er erläuterte zudem, dass derzeit sowohl Unternehmen mit digitalem Vorsprung gegenüber der Finanzverwaltung als auch umgekehrt existieren würden. Das Hauptproblem sei Transparenz und Nachvollziehbarkeit der finanzbehördlichen Datenauswertungen, insbesondere bei Insellösungen und Einzel-Systemen. Dies betreffe unter anderem die summarische Risikoprüfung und Power BI. Bei der Digitalisierung sei ein umfassender Blick auf das große Ganze notwendig. Er betonte, dass er nicht davon ausgehe, dass Betriebsprüfungen vollständig entbehrlich würden, weil neue Probleme entstünden: Während Großunternehmen häufig ein einheitliches EAP-System, etwa SAP, nutzten, arbeiteten kleinere und mittlere Unternehmen meist mit Insellösungen, deren Kompatibilität schwer herzustellen sei. Daraus entstehe eine Schnittstellenproblematik. Er betonte die Notwendigkeit eines Umdenkens in der Betriebsprüfung und hob das Vertrauen auf beiden Seiten hervor. Sich der Zukunft der Betriebsprüfung widmend lobte er § 199 Abs. 2 Sätze 2 und 3 AO als ein wertvolles Instrument, um zeitnahe Gespräche und verbindliche Meilensteine zwischen Prüfern und Unternehmen zu vereinbaren. Dies fördere ein neues Klima in der Betriebsprüfung. Zwar bezögen sich die Vorschriften auf Mitwirkungspflichten, doch könnten diese auch begrenzt werden, sodass nur die relevanten Unterlagen bereitgestellt würden, ein Ansatz, der dem fairen und angemessenen Steuervollzug diene. Er wies darauf hin, dass viele kleine und mittelständische Unternehmen bereits abgegrenzte Kontrollprozesse besitzen, etwa für Lohnsteuer, Umsatzsteuer oder Ertragssteuer, die

über Art. 97 § 38 EGAO[19] handhabbar seien. Schließlich betonte er, dass Digitalisierung ein Hilfsmittel sei, der Rechtsrahmen jedoch die Grundlage bilde und die Betriebsprüfung weiterhin ermögliche. Er hob zudem positive Ansätze in Sachsen und Brandenburg hervor, wo Strategiepapiere und Pilotprojekte die Handhabbarkeit für KMU prüften und weiterentwickelten.

Prof. Dr. *Thomas Stapperfend* richtete sich an *Lars Wargowske* und fragte, inwiefern die Überlegungen zu digitaler Betriebsprüfung auch für große Unternehmen in Brandenburg zuträfen, wie *Bettina Rodenberg* und Prof. Dr. *Christian Dorenkamp* sie dargestellt hätten. Er wollte wissen, ob er die Ansicht teile, dass für diesen Bereich künftig möglicherweise keine Betriebsprüfungen mehr nötig seien. Anschließend bat er um eine Einschätzung zur praktischen Handhabung von Art. 97 § 38 EGAO, wie von *Dirk Rose* zuvor angesprochen.

Lars Wagowske erklärte, dass die Finanzverwaltung zeitnahe Betriebsprüfungen grundsätzlich als positiv empfinde, deren Umsetzung aber oft an praktischen Schwierigkeiten scheitere. So dauere es im Durchschnitt rund 18 Monate, bis alle erforderlichen Daten von den Unternehmen vollständig bereitgestellt seien. Verzögerungen führten zu einem Stau paralleler Prüfungsfälle und dem Eindruck, es werde nichts bearbeitet, obwohl das Problem eher im fehlenden Datenfluss liege. Bei Großunternehmen, insbesondere ausländischen Inbound-Fällen, gebe es zusätzliche Herausforderungen, etwa weil andere Buchführungssysteme (HGB-Layer) erzeugt werden müssten. Das Ziel der Finanzverwaltung sei dennoch, die Prüfungen so zeitnah wie möglich durchzuführen, wobei der Wissensvorsprung der Prüfer vor Ort ein entscheidender Vorteil sei. Bezüglich Art. 97 § 38 EGAO wies er darauf hin, dass es derzeit kein offizielles Verfahren gebe. Teilweise seien Betriebsprüfer vorab konsultiert worden, um Empfehlungen zu geben und gegenseitige Abstimmungen zu ermöglichen. Er stimmte *Dirk Rose* zu, dass die Regelungen nicht nur für große Unternehmen gelten, sondern auch auf kleine und mittlere Unternehmen anwendbar seien, sofern entsprechende Mittel und Möglichkeiten zur Umsetzung gefunden würden.

Prof. Dr. *Thomas Stapperfend* bat daraufhin Prof. Dr. *Klaus-Dieter Drüen* um eine Einschätzung zur Sinnhaftigkeit von Art. 97 § 38 EGAO.

---

19 § 38 EGAO Erprobung alternativer Prüfungsmethoden

(1) Soweit im Rahmen einer Außenprüfung eines Steuerpflichtigen nach den §§ 193 bis 202 der Abgabenordnung die Wirksamkeit eines von ihm eingesetzten Steuerkontrollsystems hinsichtlich der erfassten Steuerarten oder Sachverhalte überprüft wurde und kein oder nur ein unbeachtliches steuerliches Risiko für die in § 149 Absatz 3 der Abgabenordnung genannten Steuern und gesonderten Feststellungen besteht, kann die Finanzbehörde im Benehmen mit dem Bundeszentralamt für Steuern dem Steuerpflichtigen auf Antrag unter dem Vorbehalt des Widerrufs für die nächste Außenprüfung nach § 193 Absatz 1 der Abgabenordnung Beschränkungen von Art und Umfang der Ermittlungen unter der Voraussetzung verbindlich zusagen, dass keine Änderungen der Verhältnisse eintreten. Der Steuerpflichtige hat Veränderungen des Kontrollsystems zu dokumentieren und sie der Finanzbehörde unverzüglich schriftlich oder elektronisch mitzuteilen.

(2) Ein Steuerkontrollsystem umfasst alle innerbetrieblichen Maßnahmen, die gewährleisten, dass

1. die Besteuerungsgrundlagen zutreffend aufgezeichnet und berücksichtigt werden sowie
2. die hierauf entfallenden Steuern fristgerecht und vollständig abgeführt werden.

Das Steuerkontrollsystem muss die steuerlichen Risiken laufend abbilden.

Prof. Dr. *Klaus-Dieter Drüen* bezeichnete Art. 97 § 38 EGAO als „Experimentierklausel“ und begrüßte, dass der Gesetzgeber Entwicklungen, die noch nicht vollständig abschätzbar seien, zunächst abwarte, anstatt bereits die AO zu ändern. Er stimmte der Einschätzung von *Bettina Rodenberg* zu, dass in fünf Jahren andere Themen relevant sein könnten. Spezialisierte Lösungen für digitale Unternehmen, die sich selbst überprüfen, befürwortete er, da in diesen Fällen nur noch ergänzende Prüfungen erforderlich seien. Dennoch werde die Betriebsprüfung in solchen Unternehmen nicht vollständig überflüssig. Außerdem müsse man die kleineren Unternehmen berücksichtigen, da dort die Prüfungsgewähr geringer sei. Er betonte, dass die Betriebsprüfung ein generelles Instrument sei, um die Besteuerung nach Gesetz sicherzustellen, nicht nach wirtschaftlicher Leistungsfähigkeit, und dass sie sich zwar verändern, aber nicht entfallen werde. Er unterstrich anschließend die Vorteile des Föderalismus. Unterschiedliche Ansätze der Länder, beispielsweise bei der zeitnahen Betriebsprüfung, erlaubten es, dezentral Lösungen für die digitale Realität zu entwickeln, bevor zentral einheitliche Vorgaben getroffen werden. Die dezentrale Erprobung ermögliche es, die digitale Transformation praxisnah zu gestalten. Bezüglich des Leitbilds der Vollüberprüfung stellte er klar, dass eine hundertprozentige Prüfung nie Ziel gewesen und angesichts begrenzter Ressourcen auch nicht möglich sei. Die Betriebsprüfung müsse auf alle Unternehmen verteilt werden, wobei Digitalisierung ein hilfreiches Werkzeug sei, aber kein Ersatz für menschliche Entscheidungen. Letztlich müsse die Betriebsprüfung weiterhin von Menschen durchgeführt werden, die gerichtsfeste Entscheidungen treffen. Er schloss mit der Einschätzung, dass es auch künftig Betriebsprüfungen geben werde, die sich jedoch stärker differenziert an den unterschiedlichen Unternehmensgrößen und -typen ausrichten und dabei die Digitalisierung als unterstützendes Instrument nutzbar machen.

Prof. Dr. *Thomas Stapperfend* widmete sich einem neuen Aspekt, indem er auf die Problematik der Überprüfbarkeit von Datenauswertungen hinwies. Er stellte fest, dass die Finanzverwaltung in Gerichtsverfahren Ausdrucke vorlege. Diese enthielten Schaubilder, Zahlen und Diagramme, deren Nachvollziehbarkeit er hinterfragte. Er bat Dr. *Egmont Kulosa* um eine kurze Einschätzung, welche Anforderungen an die Nachvollziehbarkeit von Datenauswertungen zu stellen seien.

Dr. *Egmont Kulosa* erklärte, dass Datenauswertungen grundsätzlich nachvollziehbar sein müssten, da Rechtsstaatlichkeit erfordere, dass Gerichte die Prüfung nachvollziehen können. Richter versuchten oft, Parallelwerte zu bereits bekannten Fällen heranzuziehen, um komplexe Berechnungen zu verstehen. Er erinnerte an Fälle aus seiner Zeit als Finanzrichter, in denen Prüfer aufwändige Berechnungen durchgeführt hatten, die selbst von den Richtern kaum nachvollzogen werden konnten. Teilweise wurden gerichtseigene Prüfer eingesetzt, andere Richter beschränkten sich darauf, nur die vom Steuerpflichtigen und dessen Berater erhobenen Einwendungen zu prüfen. Das Problem der fehlenden Nachvollziehbarkeit hätte

schon bei klassischen Schätzungsmethoden bestanden, erreiche nun aber in der digitalen Betriebsprüfung eine neue Dimension. Wenn Auswertungen nicht nachvollziehbar seien, könne eine Schätzung nicht bestätigt werden. Die Beurteilung hinge stark von den Informationen ab, die durch die „kommunizierenden Röhren“ zwischen Steuerpflichtigen, Beratern und Prüfern übermittelt werden. Zudem verfüge der BFH derzeit nur über begrenztes Material zu digitalen Prüfungsmethoden, da bisher lediglich zwei Entscheidungen (Zeitreihenentscheidung[20] und Quantilschätzungsentscheidung[21]) vorlägen.

Prof. Dr. *Thomas Stapperfend* ergänzte zustimmend, dass an den Finanzgerichten bisher nur wenige Fälle zur digitalen Betriebsprüfung eingingen. Anschließend lud er das Publikum ein, Fragen zu stellen.

C. Fragerunde

Nathalie Gossement[22] wollte wissen, welchen Einfluss die Entwicklung des internationalen Informationsaustauschs auf die Digitalisierung der Betriebsprüfung habe.

Lars Wagowske antworte, dass der automatisierte Informationsaustausch an vielen Stellen bereits genutzt werde. Eine Schwierigkeit bestehe jedoch darin, den Umgang mit Daten aus anderen Staaten zu lernen, da diese teils unter unterschiedlichen Voraussetzungen und mit unterschiedlichen Motiven erhoben würden. Er betonte, dass insbesondere die neuen Regelungen des Plattformsteuertransparenzgesetzes sowie Vorgaben zu Kryptowährungen für die Betriebsprüfung relevant seien. Es sei jedoch weiterhin eine Herausforderung, diese Daten in den deutschen Kontext zu übertragen und ein einheitliches Verständnis sicherzustellen.

D. Schlussrunde

Im Anschluss leitete Prof. Dr. *Thomas Stapperfend* die Schlussrunde ein und fragte *Dirk Rose*, ob er vor dem Hintergrund der Zukunft der Betriebsprüfung, insbesondere unter Berücksichtigung der Einschätzung, dass sich in den nächsten fünf bis zehn Jahren vieles verändern werde, seine Wünsche an die Betriebsprüfung der Zukunft benennen könne.

Dirk Rose mahnte zunächst an, dass die Wirksamkeit von Art. 97 § 38 EGAO nicht allein an der Anzahl der gestellten Anträge gemessen werden dürfe. Vielmehr sei entscheidend, dass sich die Unternehmen mit dem Instrument beschäftigten und versuchten, digitale Lösungen umzusetzen, was Zeit benötige. Er wies dann unter Bezugnahme auf Prof. Dr. *Thomas*

[20] Urteil des Bundesfinanzhofs vom 25.03.2015 – X R 20/13, DStR 2015, 1739.
[21] Beschluss des Bundesfinanzhofs vom 12.07.2017 – X B 16/17, DStRE 2017, 1316 bzw. Parallelverfahren Beschluss des Bundesfinanzhofs vom 18.7.2017 – IV B 4/17, n.v.
[22] Steuerattachée der französischen Botschaft (Berlin).

Stapperfend darauf hin, dass er nicht wolle, dass Gerichte über sämtliche Fälle entschieden, da die Unternehmen, bei denen die Prozesse funktionierten und digital gut aufgestellt seien, ohnehin nicht vor Gericht kämen. Die Fälle, die vor Gericht gelangen könnten, seien oft solche, bei denen die IT-Systeme der Unternehmen unzureichend aufgestellt seien, und er äußerte die Sorge, dass Gerichte überbordende Anforderungen stellen könnten, die praktisch nicht umsetzbar seien. Es sei sinnvoller, eine iterative Zusammenarbeit zwischen Wirtschaft und Betriebsprüfung zu fördern, um gemeinsam praxisgerechte Lösungen zu entwickeln. Er hob hervor, dass dabei ein Zielkonflikt zwischen der sicherheitsorientierten Herangehensweise der Beamten und der prozessorientierten Arbeitsweise von Unternehmen bestehe. Für die Zukunft der Betriebsprüfung wünschte er sich, dass diese grundlegend neu gedacht werde, ebenso wie der gesamte Steuervollzug, wobei er klarstellte, dass dies ein längerfristiger Prozess sei, der nicht in fünf Jahren abgeschlossen sein könne. Ein zentrales Element sei Transparenz, die Vertrauen schaffe. Er regte zudem an, über Anreizsysteme nachzudenken, um Steuerpflichtige zu mehr Offenheit zu motivieren, was gleichzeitig das Problem knapper personeller Ressourcen in der Betriebsprüfung mindern könne.

Prof. Dr. *Christian Dorenkamp* äußerte, aus seiner Sicht seien keine weiteren großen gesetzgeberischen Änderungen erforderlich, da bereits alle nötigen Instrumente vorhanden seien; entscheidend sei deren Anwendung. Er empfahl, für eine zeitnahe Betriebsprüfung auf einen einjährigen Betriebsprüfungszeitraum zu drängen, da bei längeren Zeiträumen Rückstände entstünden, was durch § 4 ABPO kaum anders möglich sei. Zudem riet er, proaktiv zu handeln und Entscheidungen zeitnah zu treffen.

Bettina Rodenberg erklärte, Digitalisierung gehe weit über die reine Prüfung von Prozessen hinaus. Sie wünsche sich, dass sowohl Unternehmen als auch die Steuerverwaltung einen „Embrace Digitalisation"-Ansatz verfolgten und betonte, dass Digitalisierung keine Angst bereiten müsse, sondern viele Herausforderungen, etwa durch den Generationenwechsel, lösen könne, wenn sie richtig eingesetzt werde. Als Beispiel verwies sie auf *ChatGPT* und verdeutlichte, dass Technologien, die vor wenigen Jahren noch nicht verfügbar waren, die Arbeitswelt in den nächsten fünf bis zehn Jahren stark verändern würden, insbesondere durch Entwicklungen wie „Generative AI"[23].

Dr. *Egmont Kulosa* führte zu Wünschen an die Zukunft aus, dass Gerichte, insbesondere die obersten Bundesgerichte, strukturell konservativ seien, unter anderem aufgrund des hohen Altersdurchschnitts der Richter. Er betonte, dass daher auch bei neuen Fällen und Methoden der Betriebsprüfung die klassischen Grundsätze wie Verhältnismäßigkeit, Vertrauensschutz,

[23] „Generative KI ist eine Teilmenge der KI, die sich auf jedes System oder jede Maschine bezieht, die menschenähnliche Aufgaben ausführen kann, indem sie ML-Modelle verwendet, um Muster in den von ihr gesammelten Daten zu identifizieren und nachzuahmen", https://www.microsoft.com/de-de/ai/ai-101/generative-ai-vs-other-types-of-ai?msockid=2cbba46971fe61113aa2b20670dc60bc (letzter Zugriff 09.10.2025).

rechtliches Gehör, Transparenz und Waffengleichheit angewendet würden. Solange die Verfassung unverändert bleibe, sei dies der Leitmaßstab und die Gerichte würden versuchen, die praktische Umsetzung rechtsstaatlich und verfassungsrechtlich einwandfrei zu prüfen.

Lars Wagowske erklärte, dass nicht alle Beteiligten von der Digitalisierung begeistert seien. Er betonte, dass insbesondere die Rücklieferung von Daten erklärbar gestaltet werden müsse, da dies zu einer Reduktion der Streitfälle führe und Diskussionen erleichtere. Dies sei nach seiner Auffassung der richtige Ansatz. Hinsichtlich der Zukunft äußerte er sich als Befürworter von „Generative AI“, wies jedoch darauf hin, dass dies ein großes Projekt sei, da sowohl die gesamte Verwaltung als auch viele Steuerpflichtige einbezogen werden müssten. Außerdem betonte er, dass Digitalisierung nicht nur Standardisierung bedeute, sondern von den Beteiligten auch akzeptiert werden müsse, da es Rückmeldungen gebe, dass manche Personen mit vollständig elektronischen Vorgängen unzufrieden seien.

Prof. Dr. *Klaus-Dieter Drüen* äußerte, dass er sich in seinen Wünschen zur Zukunft der Betriebsprüfung der Forderung von Prof. Dr. *Christian Dorenkamp* nach Zeitnähe anschließe, da dies eine elementare Anforderung sei, die auch im Zusammenhang mit der Verkürzung der Festsetzungsfristen stehe. Er betonte, dass bei der heutigen Diskussion die Betriebsprüfung als intensivste Form der Verifikation im Vordergrund stand, nicht das materielle Recht. Dazu brachte er die Überlegung ein, dass weniger leistungsfähigkeitsgerechte Besteuerung und mehr Typisierung eine praktikable Alternative sein könnten, etwa durch Schätzung von Verbrauch pro Kopf.

Prof. Dr. *Thomas Stapperfend* dankte den Teilnehmenden und wies auf das digitale 98. Steuergespräch am 1. Dezember zum Thema „Steuerliche Arbeitsanreize | Insbesondere Aktivrente, steuerfreie Überstundenzuschläge und Mehrarbeitsprämien“ hin.

Impressum

Berliner Steuergespräche e.V.

c/o POELLATH

Potsdamer Platz 5

10785 Berlin

www.berlinersteuergespraeche.de

berliner.steuergespraeche@pplaw.info