

Darf es ein bisschen mehr sein?

Roger Odenthal, Köln

Dem Charme dieser harmlosen Frage folgten zahlreiche interessierte Zuhörer auf einer Münsteraner Brennpunkt-Veranstaltung zu dem Thema „Schätzungen in der Betriebsprüfung“. Anlass war eine aktuelle und aufsehenerregende Entscheidung des Bundesfinanzhofes, die von BFH-Richter Prof. Dr. Georg Nöker mit ihren zugrunde liegenden Aspekten umfassend erläutert wurde. Das Urteil folgte einem Streit über die mehrjährige Gewinnermittlung (2001 bis 2003) eines Restaurantbetriebes. Angesichts zahlreicher *formeller Mängel* der Kassenbuchführung:

- lückenhafte Aufbewahrung von Tagesendsummen-Bons
- keine Programmierprotokolle zur elektronischen Registrierkasse
- fehlende Kassenberichte für eine hiervon getrennte (Bar-)Thekenkasse
- Warenendbestände, die sich nicht auf eine Inventuraufnahme stützen konnten

sowie weiteren *materiellen Unrichtigkeiten* der Rechnungslegung wurden dessen Besteuerungsgrundlagen durch das zuständige Finanzamt mit Hilfe eines Zeitreihenvergleichs geschätzt. Insbesondere die Schätzungsmethode war angesichts der sich hieraus ergebenden Höhe Gegenstand der höchstrichterlichen Würdigung.

Die Meinung der Richter

Die mit Hilfe statistischer Berechnungen (Zeitreihenvergleich) seitens des Finanzamtes unterstellten Rohgewinnaufschläge können nach der BFH-Entscheidung ein sich hieraus ergebendes, wenig plausibles Schätzungsergebnis nicht ausreichend begründen. Über den betrachteten Einzelfall hinaus sind die Risiken statistischer Schätzverfahren dabei nach Meinung des Gerichts methodenimmanent. So rücken betriebliche Besonderheiten z.B. bei Einkauf, Lagerung oder Preisgestaltung gegenüber allgemeinen, an statistischen Durchschnittswerten orientierten Zahlenentwicklungen in den Hintergrund. Das Risiko fehlerhaft konfekzionierter Hochrechnungsverfahren mit Schätzungsergebnissen, die sich denkllogisch kaum mehr nachvollziehen lassen, ist daher besonders hoch. Hierauf stützt der Bundesfinanzhof seine Einschätzung, dass der *Zeitreihenvergleich* primär als mathematisch-statistische *Verprobungsmethode* und somit als *prüferisches Handwerkszeug*, nicht jedoch als *hinreichend zuverlässige Schätzungsmethode* einzuordnen sei. Zu letzterem fehle es an geeigneten empirischen und gutachterlichen Stellungnahmen von Statistikern, welche gegebenenfalls zu einer nach § 96 Abs. 1 Satz 1 FGO erforderlichen Sicherheit der Schätzung beitragen könnten.

Damit ist der Zeitreihenvergleich als Schätzmethode nicht grundsätzlich aus der Welt, er muss sich jedoch verschärften Einsatzvoraussetzungen stellen, die der Bundesfinanzhof in einer „Drei-Stufen-Theorie“ aufführt:

- Stufe 1: Formal ordnungsgemäße Buchführung (bis auf geringe Mängel)
Unter den vorgenannten Randbedingungen ist der Zeitreihenvergleich aus Sicht des Gerichts alleine nicht geeignet, um die materielle Unrichtigkeit einer Buchführung nachzuweisen.
- Stufe 2: Formal nicht ordnungsgemäße Buchführung ohne das deren materielle Unrichtigkeit konkret nachgewiesen ist
Soweit der Steuerpflichtige hier im Rahmen seiner Möglichkeiten an der Aufklärung unklarer Sachverhalte mitwirkt und nicht ohnehin komplexe Buchführungsverhältnisse vorherrschen, sind analytische Schätzmethoden, wie z.B. Vermögenszuwachs- und Geldverkehrsrechnung sowie Aufschlagskalkulation vorrangig anzuwenden.

- Stufe 3: Eine Buchführung weist neben formellen auch substantielle materielle Mängel auf
Soweit sich bei offensichtlichen Unrichtigkeiten z.B. nicht gebuchten Wareneinkäufen, un versteuerten Betriebseinnahmen oder rechnerischen Kassenfehlbeträgen nicht andere analytische Schätzungsmethoden aufdrängen, kann die Steuerverwaltung auf einen technisch korrekt durchgeführten Zeitreihenvergleich zurückgreifen

Angesichts der mit statistischen Hochrechnungen verbundenen Handhabungs- und Irrtumsrisiken ordnet der Bundesgerichtshof den Zeitreihenvergleich damit erkennbar als Ultima Ratio hinter andere, schlüssigere und leicht nachvollziehbare Schätzverfahren ein. Gleichzeitig wurde seitens Prof. Nöcker jedoch ausdrücklich darauf hingewiesen, dass den Finanzgerichten im Einzelfall eigene Schätzbefugnisse zuständen und Steuerpflichtige, die Mitwirkungspflichten bei Sachverhaltsaufklärungen erkennbar nicht nachkämen, nachteilige Schätzungen billigend in Kauf nähmen. Hiervon betroffenen Steuerpflichtigen seien dann allerdings ausnahmslos alle Aspekte der fiskalischen Schätzung, neben dem vollständigen Zahlenwerk, zugehörige Excel-Tabellen einschließlich rechentechnischer Verknüpfungen und Berechnungsmakros für eine nachvollziehbare Überprüfung zur Verfügung zu stellen.

Eine den richterlichen Vortrag abschließende Bemerkung galt dem *statistischen Prüfungskonzept „Summarische Risikoprüfung (SRP)“* der Finanzverwaltung, welches der steuerlichen Betriebsprüfung in einigen Bundesländern zunehmend zugrunde liegt. Wie dieses im Kontext der höchstrichterlichen Entscheidung einzuordnen ist, bleibt zunächst offen, da, so der Eindruck, diese Methode wohl in erster Linie der Überprüfung von Auffälligkeiten (Verprobung) und nicht der Schätzung zugerechnet wird.

Die Sicht der Finanzverwaltung

Es ist unmittelbar einsichtig, dass die Finanzverwaltung eine aus ihrer Sicht vorteilhafte und mit viel Aufwand etablierte Schätzmethode nur ungerne aus der Hand geben möchte. So waren denn die Ausführungen von Herrn Landesregierungsdirktor Arno Becker (Referatsleiter für Außenprüfungsdienste der Oberfinanzdirektion NRW) erkennbar darauf abgestimmt, dem einschränkenden BFH-Urteil die langjährige Erfahrung und Reife bei der Anwendung statistischer Schätzmethoden durch die Finanzämter entgegnen zu lassen.

So stimme man mit dem Bundesfinanzhof grundsätzlich darin überein, dass Verprobungs- und Schätzverfahren klar voneinander abzugrenzen seien. Weiterhin sei es selbstverständlich, dass von einer Schätzung betroffenen Steuerpflichtigen eine ausreichende sowie vollständige Dokumentation aller Schätzungsgrundlagen für eigene Nachprüfungen zur Verfügung gestellt werden müsse. Diese habe gleichermaßen Anspruch auf eine handwerklich und technisch korrekte Ausführung bei der Anwendung des Schätzverfahrens. Darüber hinaus gelte es jedoch zu bedenken, dass eine über Jahre optimierte Herangehensweise der Steuerbehörden den statistischen Zeitreihenvergleich soweit verbessert habe, dass er heute den analytischen Schätzmethoden, wie z.B. der Aufschlag-Ausbeutekalkulation oder Geldverkehrsrechnung zumindest ebenbürtig wenn nicht sogar überlegen sei.

Innerhalb des für Schätzungen entwickelten Verwaltungsprogramms „FinANZ“ erfolge eine angemessene Behandlung unplausibler Extremwerte oder Ausreißer in Zeitreihenvergleichen mit Hilfe zusätzlicher Sensitivitätsanalysen, Ausbeutekalkulationen oder Verteilungsbetrachtungen z.B. zu Speisen und Getränken. Die darüber hinaus für Verprobungszwecke eingesetzte Software zur summarischen Risikoprüfung (SRP) lege zusätzlich mittels Massendatenanalysen ein systematisches und vorlagengestütztes Verprobungsnetz über die Buchhaltungsdaten eines Steuerpflichtigen. Dieses ermögliche den Abgleich großer betrieblicher Datenmengen und hierdurch die Gewinnung eines Überblicks über operative Verhältnisse geschäftlicher Aktivitäten sowie Schlüssigkeitsprüfungen von Waren- oder Leistungsflüssen, Entgeltstrennungen sowie vergleichbaren Faktoren. Hierbei kämen sowohl analytische Prüfungstechniken z.B. betriebswirtschaftliche Leistungsanalysen als auch mathematisch-statistische Wahrscheinlichkeitsbetrachtungen in Form digitaler Ziffernzusammenstellungen (Benford-Muster) oder Log-Norm-Verteilungen zur Anwendung. Der jeweils agierende Betriebsprüfer werde durch das Programm geführt und mittels Warnhinweise auf eine unsachgemäße Anwendung aufmerksam gemacht. Darüber hinaus seien die Analyseergebnisse reproduzierbar und ausreichend dokumentiert.

Die Erläuterungen zur Anwendung des Zeitreihenvergleichs als bevorzugter Schätzmethode wurden von einer eindrucksvollen Demonstration verschiedenster Manipulationssoftware für elektronische Kassensysteme begleitet. Vor diesem Hintergrund stelle sich den Steuerbehörden bei auffälligen Datenkonstellationen stets die Frage, welche Alternativnachweise, neben statistischen Hochrechnungen, vermuteten, spurlosen Umsatzunterdrückungen entgegengestellt werden könnten? Diesem Zusammenhang galt der nachdrückliche Hinweis auf eine strenge Aufbewahrungspflicht der Programmierprotokolle von Kassensystemen. Alleine deren Fehlen stelle bereits einen *schwerwiegenden formellen Mangel* des Buchführungsverfahrens und damit einen ersten Schritt in die Schätzung dar. Hiervor könne sich der Steuerpflichtige aus Sicht der Finanzverwaltung gegebenenfalls dadurch schützen, dass er zur Aufzeichnung von Barkassengeschäften technisch manipulationsgeschützte Kassensicherungssysteme, z.B. nach dem „INSIKA-Konzept“ einsetze.

Worauf dürfen sich die Steuerpflichtigen einstellen?

Die kontroverse Erörterung hinterlässt den Eindruck, dass die Finanzverwaltung an der *Schätzmethode* „Zeitreihenvergleich“ mit Hinweis auf sich fortentwickelnde Anwendungserfahrungen auch in der Zukunft festhalten wird. Bedenklich erscheint in diesem Zusammenhang die zusätzlich begleitende Anwendung statistischer Verprobungstechniken (Summarische Risikoprüfung) auf der Grundlage vermuteter Ziffern- und logarithmierter Werteverteilungen. Auf deren zahlreiche Unzulänglichkeiten einschließlich hieraus resultierender unzutreffender Verdächtigungen, denen sich ein Steuerpflichtiger ohne mathematischen Sachverstand kaum entgegenzustellen vermag, soll an dieser Stelle noch einmal ausdrücklich hingewiesen werden. Soweit sich hieran eine gleichermaßen auf statistische Zusammenhänge gründende Schätzung (Zeitreihenvergleich) anschließt, kann die Gefahr eines sich selbst befördernden Zirkelschlusses nicht ausgeschlossen werden.

Dass die Probleme einer unsachgemäßen Heranziehung von Schätzparametern bei statistischen Verfahren durch örtliche Betriebsprüfer entgegen den Ausführungen der Finanzverwaltung weiterhin nicht gebannt sind, zeigt ein aktuelles Urteil des Finanzgerichts Münster. Der zugrunde liegende Sachverhalt reicht bis in das Jahr 2011. Auch hier führten bei unstreitig formellen Mängeln der Buchführung nachfolgende statistische Ergebnisschätzungen der Steuerverwaltung zu nicht plausiblen Werten. In der Urteilsbegründung weist das Gericht die Steuerbehörden ausdrücklich an, im Rahmen ihres Ermessens *„die Schätzmethode so zu wählen, dass die Besteuerungsgrundlagen durch Wahrscheinlichkeitsüberlegungen der Wirklichkeit möglichst nahekommen. Schätzungen müssen insgesamt in sich schlüssig sein. Ihre Ergebnisse müssen darüber hinaus wirtschaftlich vernünftig und möglich sein.“* Insbesondere, wenn lediglich formelle Mängel zu verzeichnen sind, der Nachweis materieller Unrichtigkeiten aber nicht geführt werden kann, seien *„Schätzungsmethoden, die auf betriebsinternen Daten aufbauen oder in andere Weise die individuellen Verhältnisse des jeweiligen Steuerpflichtigen berücksichtigen (z.B. eine Aufschlagskalkulation) grundsätzlich vorrangig heranzuziehen.“*

Die betriebliche Steuerprüfung bleibt insoweit angehalten, ihre Arbeit sorgfältig sowie mit Blick auf allfällige betriebliche oder geschäftliche Besonderheiten durchzuführen. Der Steuerpflichtige sowie seine Berater erhalten durch das vorgestellte Drei-Stufen-Modell des BFH einen Orientierungsrahmen, innerhalb dessen sich Ergebnisschätzungen steuerlicher Betriebsprüfer zuverlässiger einordnen lassen. Gleichzeitig wurde deutlich, dass sich ein gewichtiger Teil hiermit in Verbindung stehender Probleme elegant dadurch entschärfen lässt, dass zwingend manipulationssichere Kassensysteme eingesetzt werden müssen. Hier ist der Gesetzgeber gefordert, nicht länger den kaum nachvollziehbaren Einwänden von Interessengruppen zu folgen, sondern im Interesse der Steuerehrlichkeit eine Regelung vorzuschreiben, wie sie in Nachbarländern längst übliche und unproblematische Praxis ist.