Urteil vom 25. März 2015, X R 20/13

Anforderungen an die Schätzung mittels eines Zeitreihenvergleichs

BFH X. Senat

AO § 147 Abs 1 Nr 1, AO § 158, AO § 162, AO § 5, FGO § 96 Abs 2, EStG § 4 Abs 1, EStG § 4 Abs 4, FGO § 76 Abs 2, AO § 147 Abs 1 Nr 5

vorgehend FG Münster, 25. Juli 2012, Az: 4 K 2071/09 E,U

Leitsätze

1. Die Durchführung eines Zeitreihenvergleichs setzt voraus, dass im Betrieb das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum weitgehend konstant ist. Es darf zudem im maßgebenden Zeitraum nicht zu solchen Änderungen in der Betriebsstruktur gekommen sein, die --nicht anderweitig behebbare-- wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringen.

2. Bei einer Buchführung, die formell ordnungsgemäß ist oder nur geringfügige formelle Mängel aufweist, kann der Nachweis der materiellen Unrichtigkeit grundsätzlich nicht allein aufgrund der Ergebnisse eines Zeitreihenvergleichs geführt werden.

3. Ist die Buchführung formell nicht ordnungsgemäß, sind aber materielle Unrichtigkeiten der Einnahmenerfassung nicht konkret nachgewiesen, können die Ergebnisse eines Zeitreihenvergleichs nur dann einen Anhaltspunkt für die Höhe der erforderlichen Hinzuschätzung bilden, wenn andere Schätzungsmethoden, die auf betriebsinternen Daten aufbauen oder in anderer Weise die individuellen Verhältnisse des jeweiligen Steuerpflichtigen berücksichtigen, nicht sinnvoll einsetzbar sind. Bei verbleibenden Zweifeln können Abschläge in einem Umfang geboten sein, der über eine bloße Abrundung hinausgeht.

4. Steht bereits aus anderen Gründen fest, dass die Buchführung sowohl formell als auch materiell unrichtig ist und übersteigt die nachgewiesene materielle Unrichtigkeit eine von den Umständen des Einzelfalls abhängige Bagatellschwelle, können die Ergebnisse eines --technisch korrekt durchgeführten-- Zeitreihenvergleichs auch für die Ermittlung der erforderlichen Hinzuschätzung der Höhe nach herangezogen werden, sofern sich im Einzelfall keine andere Schätzungsmethode aufdrängt, die tendenziell zu genaueren Ergebnissen führt und mit vertretbarem Aufwand einsetzbar ist.

5. Bei einem programmierbaren Kassensystem stellt das Fehlen der aufbewahrungspflichtigen Betriebsanleitung sowie der Protokolle nachträglicher Programmänderungen einen formellen Mangel dar, dessen Bedeutung dem Fehlen von Tagesendsummenbons bei einer Registrierkasse oder dem Fehlen von Kassenberichten bei einer offenen Ladenkasse gleichsteht und der daher grundsätzlich schon für sich genommen zu einer Hinzuschätzung berechtigt.

Tenor

Auf die Revision der Kläger wird das Urteil des Finanzgerichts Münster vom 26. Juli 2012 4 K 2071/09 E,U aufgehoben.

Die Sache wird an das Finanzgericht Münster zurückverwiesen.

Diesem wird die Entscheidung über die Kosten des Revisionsverfahrens übertragen.

Tatbestand

I

..

1 Die Kläger und Revisionskläger (Kläger) sind Eheleute, die in den verbliebenen Streitjahren 2001 und 2003 zur Einkommensteuer zusammenveranlagt wurden. Der Kläger betrieb von 1996 bis zum Frühjahr 2002 und erneut ab dem Frühjahr 2003 eine Schank- und Speisewirtschaft in Räumen, die er von einer Brauerei angemietet hatte. Die Klägerin war im Betrieb angestellt.

2 Der Kläger ermittelte seinen Gewinn durch Betriebsvermögensvergleich. Den größten Teil seiner Bareinnahmen rechnete er über eine elektronische Registrierkasse ab. Daneben führte er für die Thekeneinnahmen eine von der Registrierkasse getrennte Barkasse.

3 Im Rahmen einer Außenprüfung beanstandete der Prüfer die Ordnungsmäßigkeit der Kassenführung und stützte sich hierbei auf die folgenden Punkte:

- Zwar ist für jeden Öffnungstag des Jahres 2001 ein Tagesendsummenbon vorhanden; die fortlaufende Nummerierung dieser Bons weist für die Zeit vom 1. Januar bis 16. März 2001 aber Unterbrechungen auf. Die Bons mit den dazwischenliegenden Nummern haben die Kläger nicht vorgelegt und hierzu vorgetragen, es handele sich um "Leerbons".

- Für das Jahr 2003 liegen zwar fortlaufend nummerierte Tagesendsummenbons vor. Allerdings ist auf diesen Bons kein Datum ausgedruckt.

- Die Kläger haben weder Programmierprotokolle der Registrierkasse noch die Speisekarte für das Jahr 2001 vorgelegt.

- Ausweislich der Feststellungen des Finanzgerichts (FG) --im Betriebsprüfungsbericht ist dieser Umstand hingegen nicht erwähnt-- wurden für die Thekenkasse keine Kassenberichte erstellt, in denen die sich anhand der Aufzeichnungen ergebenden Soll-Kassenbestände mit den durch Auszählen ermittelten Ist-Beständen abzugleichen gewesen wären. Vielmehr wurden die täglichen Bestände dieser Kasse lediglich rechnerisch durch Gegenüberstellung der aufgezeichneten Einnahmen und Ausgaben ermittelt.

- Der Kläger hat die Warenendbestände zum Ende der Streitjahre nicht durch Inventuren, sondern im Wege der Schätzung ermittelt.

- Für das Jahr 2003 hat der Prüfer vier Bareinlagen in Höhe von insgesamt 4.164,78 € als ungeklärt angesehen. Hierzu hat das FG allerdings keine Feststellungen getroffen.

4 Der Prüfer erhöhte die vom Kläger erklärten Erlöse um Hinzuschätzungen, für deren Höhe er sich auf einen sog. "Zeitreihenvergleich" stützte. Dabei ging er wie folgt vor:

- Aus den Eingangsrechnungen des Klägers ermittelte der Prüfer den Wareneinkauf für jede Kalenderwoche des Prüfungszeitraums. Nach dem --von den Klägern teilweise bestrittenen-- Vortrag des Beklagten und Revisionsbeklagten (Finanzamt --FA--) habe der Prüfer hierbei das Lieferdatum zugrunde gelegt, sofern ein solches in den Eingangsrechnungen angegeben gewesen sei. Ansonsten habe er das Rechnungsdatum herangezogen. Sofern der Kläger beim jeweiligen Lieferanten weniger als einmal wöchentlich eingekauft habe, habe der Prüfer den Wert des Einkaufs gleichmäßig auf den Zeitraum zwischen den aufeinanderfolgenden Einkäufen verteilt.

- Den Wareneinkauf minderte der Prüfer um die sich aus der Buchführung des Klägers --auf der Grundlage der von der Finanzverwaltung hierfür veröffentlichten Pauschbeträge-- ergebenden Werte für Sachentnahmen der Familie der Kläger und für die Beköstigung des Personals. Diese Beträge verteilte er gleichmäßig auf die Kalenderwochen eines Jahres.

- Für das Jahr 2003 minderte der Prüfer den Wareneinkauf ferner um den vom Kläger geschätzten Warenendbestand zum 31. Dezember 2003 (10.800 €), der ebenfalls gleichmäßig auf die einzelnen Kalenderwochen verteilt wurde.

- Den sich danach ergebenden wöchentlichen Wareneinsatz verglich der Prüfer mit den vom Kläger für die jeweilige Kalenderwoche erklärten Einnahmen. Aus diesem Vergleich ergab sich für jede Woche ein Rohgewinnaufschlagsatz.

- Anschließend ermittelte der Prüfer für jede Kalenderwoche den Durchschnittswert aus den Rohgewinnaufschlagsätzen dieser Woche und der neun vorangehenden Wochen.

- Seiner Schätzung legte der Prüfer den höchsten Zehn-Wochen-Durchschnittswert zugrunde, der sich im jeweiligen Kalenderjahr ergab. Dies waren für das Jahr 2001 die Kalenderwochen 23 bis 32 mit einem Rohgewinnaufschlagsatz (hier: Vergleich der Brutto-Einnahmen mit dem Netto-Wareneinsatz) von 241,54 %, für das Jahr 2003 die Kalenderwochen 28 bis 37 mit einem Brutto/Netto-Rohgewinnaufschlagsatz von 263,42 %.

- Diese Rohgewinnaufschlagsätze wendete der Prüfer für das gesamte jeweilige Kalenderjahr auf den tatsächlich vom Kläger erklärten Wareneinsatz an.

5 Auf dieser Grundlage ermittelte der Prüfer die folgenden Hinzuschätzungsbeträge:

Jahr	2001	2003
Wareneinsatz netto	200.244,83 DM	73.368,42 €
höchster RAS brutto/netto	241,54 %	263,42 %
zum Vergleich: RAS lt. Kläger	228,15 %	201,09 %
Erlöse brutto lt. Prüfer	683.908,83 DM	266.634,31 €
Erlöse brutto lt. Kläger	657.095,38 DM	220.905,86 €
Mehrerlöse brutto lt. Prüfer	26.813,44 DM	45.728,46 €
Mehrerlöse netto lt. Prüfer	23.115,03 DM	39.421,08 €
Hinzuschätzung netto lt. Prüfer (nach Abrundung auf den nächsten durch 5.000 teilbaren Betrag)	20.000,00 DM	35.000,00 €

6 Für das --im Revisionsverfahren nicht mehr streitbefangene-- Jahr 2002 nahm der Prüfer wegen aus seiner Sicht ungeklärt gebliebener Vermögenszuwächse eine Hinzuschätzung von 40.000 € netto vor.

7 Das FA erließ entsprechend geänderte Einkommen- und Umsatzsteuerbescheide für die Streitjahre. Im Einspruchsverfahren legten die Kläger Geldverkehrsrechnungen für die Jahre 2001 und 2002 samt Unterlagen zu ihren außerbetrieblichen Einnahmen sowie zu Vermögensumschichtungen vor. Für das Jahr 2003 beantragten sie, weitere Schuldzinsen als Betriebsausgaben zu berücksichtigen. Gegen den Zeitreihenvergleich wandten sie sich mit dem folgenden Vorbringen:

- Für die Durchführung des Zeitreihenvergleichs sei die exakte Zuordnung des wöchentlichen Wareneinsatzes zu den Erlösen der jeweiligen Woche von herausragender Bedeutung. Der Schluss vom Waren*einkauf* --nur insoweit bestehe eine gesetzliche Aufzeichnungspflicht-- auf den wöchentlichen Waren*einsatz* sei aber wegen des Fehlens einer Verpflichtung zu wöchentlichen Inventuren mit großen Unsicherheiten belastet. Zwar werde durch die Bildung eines Durchschnittswerts über einen Zehn-Wochen-Zeitraum eine gewisse Nivellierung vorgenommen; zufällige Verschiebungen am Anfang oder Ende eines solchen Zehn-Wochen-Zeitraums könnten aber immer noch erhebliche rechnerische Auswirkungen auf den sich ergebenden Rohgewinnaufschlagsatz haben.

- Bereits die betriebs- und saisonüblichen Schwankungen in der Lagerhaltung sowie hinsichtlich der Ein- und Verkaufspreise und der Verteilung zwischen den Warenarten würden aufgrund der Systematik des Zeitreihenvergleichs zwingend zu rechnerischen Mehrergebnissen führen.

- Es entspreche nicht dem tatsächlichen Ablauf, den Aufbau des zum 31. Dezember 2003 vorhandenen Warenendbestands gleichmäßig auf alle Kalenderwochen des Jahres 2003 zu verteilen.

- Am letzten Augustwochenende (34. Kalenderwoche 2001 bzw. 35. Kalenderwoche 2003) finde jeweils das X-Fest statt, das zwar zu erheblich überdurchschnittlichen Wochenumsätzen, nicht aber zu einem entsprechenden Anstieg des Wareneinkaufs in derselben Kalenderwoche führe.

- Im Jahr 2003 sei ein gerade kostendeckender Mittagstisch angeboten worden.

- Der Warenverderb sei unberücksichtigt geblieben.

8 In seinen Einspruchsentscheidungen setzte das FA die Hinzuschätzung für 2002 auf 15.000 € herab und wies die Einsprüche im Übrigen zurück. Im Klageverfahren legten die Kläger auch für 2003 eine Geldverkehrsrechnung vor und erklärten aus vorhandenen Geldanlagen Zinseinnahmen für 2001 nach, die allerdings unterhalb der damals geltenden Freibeträge lagen.

9 Das FG gab der Klage für 2002 in vollem Umfang statt und berücksichtigte für 2003 zusätzliche Schuldzinsen in Höhe von 3.764 € als Betriebsausgaben. Im Übrigen wies es die Klage ab (Entscheidungen der Finanzgerichte --EFG-- 2012, 1982). Zur Begründung des klageabweisenden Teils seiner Entscheidung führte es aus, die Buchführung des Klägers sei in den Jahren 2001 und 2003 formell nicht ordnungsgemäß gewesen. Diese formellen Mängel seien wegen der entscheidenden Bedeutung der Bareinnahmen für den Betrieb des Klägers auch Anlass, die sachliche Richtigkeit der Buchführung anzuzweifeln, da es wegen der nur rechnerischen Führung der Kasse (gemeint wohl: Thekenkasse) an der Kassensturzfähigkeit fehle.

10 Der Zeitreihenvergleich stelle im Streitfall eine geeignete Schätzungsmethode dar. Die Wahl der Schätzungsmethode stehe im Ermessen des FA bzw. des FG; Unschärfen gingen zu Lasten des Steuerpflichtigen. Die von der instanzgerichtlichen Rechtsprechung anerkannte Methode des Zeitreihenvergleichs basiere darauf, dass es in der Praxis kaum möglich sei, über kurze Zeiträume genau denjenigen Wareneinkauf zu verschweigen, mit dem nicht verbuchte Erlöse erzielt würden. Im Betrieb des Klägers habe es in den Streitjahren keine wesentlichen Änderungen in der Betriebsführung oder -struktur gegeben, die einem Zeitreihenvergleich entgegenstehen könnten. Saisonale oder jahreszeitliche Schwankungen des Geschäfts seien nicht erkennbar, weil die wochenweise ermittelten Rohgewinnaufschlagsätze nicht in Abhängigkeit von der Jahreszeit, sondern auch innerhalb einer Jahreszeit unterschiedlich seien. Das X-Fest könne sich nicht ausgewirkt haben, da es nur im Jahr 2003, nicht aber im Jahr 2001 in den vom FA herangezogenen Zehn-Wochen-Zeitraum falle, und auch für das Jahr 2003 der höchste Rohgewinnaufschlagsatz des Zehn-Wochen-Zeitraums gerade nicht in die Woche des X-Fests falle. Der Warenendbestand 2003 sei zu Recht auf das gesamte Jahr verteilt worden, da davon auszugehen sei, dass er während des gesamten Jahres aufgebaut worden sei. Für die Kläger sei dies günstiger als eine Berücksichtigung in einem kurzen Zeitraum, die dann in diesem Zeitraum zu überproportional hohen rechnerischen Wareneinsätzen führen würde. Die vom FA vorgenommene Abrundung gleiche weitere Unschärfen der Schätzung aus. Das Schätzungsergebnis bewege sich im mittleren Bereich der Richtsätze für Gast-, Speise- und Schankwirtschaften (Netto/Netto-Mittelwert: 213 % bei einer Spanne von 150 bis 317 % für 2001 bzw. 144 bis 317 % für 2003).

11 Die von den Klägern eingereichten Geldverkehrsrechnungen seien als Schätzungsgrundlage nicht geeignet. Ihnen seien die sich aus der Buchführung ergebenden Betriebseinnahmen zugrunde gelegt worden, obwohl die

Buchführung nicht ordnungsgemäß sei. Für 2001 komme hinzu, dass die Auszahlung aus einem Lebensversicherungsvertrag doppelt angesetzt und die Herkunft der aus dem Ausland stammenden Mittel nicht nachgewiesen worden sei. Ohnehin bestehe kein Anspruch auf die Anwendung einer bestimmten Schätzungsmethode.

12 Für das Jahr 2002 sei die Klage hingegen begründet. Dem FA stehe insoweit keine Schätzungsbefugnis zu, weil die Buchführung des Klägers keine Mängel aufgewiesen habe. Es seien auch keine ungeklärten Vermögenszuwächse zu verzeichnen, weil die vom FA vorgenommene Geldverkehrsrechnung methodische Mängel aufweise.

13 Mit ihrer Revision wiederholen und vertiefen die Kläger ihre Einwendungen gegen den vom FA durchgeführten Zeitreihenvergleich. Ergänzend führen sie die folgenden Gesichtspunkte an:

- Selbst wenn man davon ausgehe, dass die wochenweisen Schwankungen im Einkauf frischer Lebensmittel sich über einen Zehn-Wochen-Zeitraum ausgleichen würden, gelte dies nicht hinsichtlich alkoholischer Getränke, die auf Vorrat gekauft würden.

- Dem von einem Zeitreihenvergleich betroffenen Steuerpflichtigen erschließe sich weder der Ablauf noch das Ergebnis dieser Schätzungsmethode. Dies kollidiere sowohl mit dem für Verwaltungsakte geltenden Begründungserfordernis (§ 121 der Abgabenordnung --AO--) als auch mit dem Anspruch auf effektiven Rechtsschutz (Art. 19 Abs. 4 des Grundgesetzes --GG--). Es sei dem Steuerpflichtigen nicht möglich, auf die methodischen Mängel des Zeitreihenvergleichs zu reagieren, weil hierfür im Betrieb keine Daten erhoben würden und auch nicht erhoben werden müssten.

- Die Anwendung des höchsten sich im jeweiligen Kalenderjahr ergebenden Zehn-Wochen-Rohgewinnaufschlagsatzes widerspreche den Denkgesetzen. Zwar könnten Unterschiede zwischen den wöchentlichen Aufschlagsätzen auf Manipulationen hindeuten; sie könnten aber ebenso auf den methodischen Mängeln des Zeitreihenvergleichs --insbesondere auf betriebsüblichen Schwankungen-- beruhen.

- Der Biereinkauf sei gerade in den Wochen ab Ende November 2003 sehr hoch gewesen, so dass ein überproportionaler Teil des Warenendbestands in dieser Zeit aufgebaut worden sei.

- Während der Außenprüfung sei beim Kläger nicht der geringste Anhaltspunkt für Schwarzeinkäufe gefunden worden.

14 Die Kläger haben im Revisionsverfahren ausdrücklich zugestanden, dass ihre Buchführung formell nicht ordnungsgemäß war.

15 Nach Ergehen des angefochtenen Urteils hat das FA am 20. Dezember 2012 und am 4. Februar 2013 geänderte Einkommensteuerbescheide für 2001 erlassen. Der Streitstoff des vorliegenden Verfahrens wird dadurch nicht berührt, weil die Besteuerungsgrundlagen im Bescheid vom 4. Februar 2013 denen entsprechen, die bereits in dem von den Klägern ursprünglich angefochtenen Änderungsbescheid angesetzt worden sind.

16 Die Kläger beantragen sinngemäß,
das angefochtene Urteil aufzuheben, soweit darin die Klage abgewiesen worden ist, und den Einkommensteuerbescheid 2003 sowie die Umsatzsteuerbescheide 2001 und 2003, jeweils vom 17. März 2008 und unter Aufhebung der Einspruchsentscheidungen vom 15. Mai 2009, sowie den Einkommensteuerbescheid 2001 vom 4. Februar 2013 dahingehend zu ändern, dass die Hinzuschätzungen zu den Betriebseinnahmen bzw. umsatzsteuerpflichtigen Entgelten in Höhe von 20.000 DM (2001) bzw. 35.000 € (2003) --hinsichtlich der Einkommensteuer unter gegenläufiger Kürzung der Gewerbesteuer-Rückstellung-- rückgängig gemacht werden.

17 Das FA beantragt,
die Revision zurückzuweisen.

18 Es trägt vor, die Methode des Zeitreihenvergleichs sei nicht neu, sondern werde von Wirtschaftsprüfern seit Jahrzehnten zur betriebswirtschaftlichen Analyse eingesetzt. Für die Außenprüfung habe sie in den letzten Jahren durch das Datenzugriffsrecht und spezielle Software, die für die Finanzverwaltung entwickelt worden sei, an Bedeutung gewonnen. Sie beruhe darauf, dass sich selbst bei einem Verschweigen sowohl von Wareneinkäufen als

auch von Erlösen im Wochenvergleich Auffälligkeiten ergeben würden. Die Methode werde vorzugsweise in Branchen mit eher geringer Lagerhaltung angewendet, bei denen sich die Umsätze relativ gleichmäßig auf das Jahr verteilten (z.B. Gastronomie). Ihr liege die Annahme zugrunde, dass es nicht sinnvoll wäre, bei jederzeit verfügbaren und leicht verderblichen Waren eine nennenswerte Lagerhaltung zu betreiben. Wenn diese Methode zum Ausweis größerer Unterschiede zwischen den wöchentlichen Rohgewinnaufschlagsätzen führe und sich derartige Unterschiede auch nach Prüfung der tatsächlichen Verhältnisse nicht mit den Besonderheiten des jeweiligen Betriebs erklären ließen, werde vermutet, dass Verkäufe und/oder Einkäufe nicht zutreffend in der Buchführung erfasst worden seien. Weil der Zeitreihenvergleich auf betrieblichen Daten beruhe, sei er einer Richtsatzschätzung grundsätzlich überlegen. Er lasse kaum Raum für Zweifel an der unzutreffenden Verbuchung der Erlöse. Grundlage der Hinzuschätzungen seien aber auch in diesem Fall nicht Schwarzeinkäufe, sondern die fehlende Ordnungsmäßigkeit der Buchführung.

19 Die Kläger hätten sich auf die Aufzählung möglicher Unsicherheiten dieser Schätzungsmethode beschränkt, aber nicht hinreichend dargelegt, aus welchen Gründen der bei ihnen konkret durchgeführte Zeitreihenvergleich zu falschen Ergebnissen geführt habe. Die Schätzung sei hinreichend begründet worden; alle Berechnungen seien den Klägern zur Verfügung gestellt worden. Die Verteilung des Warenendbestands sei Gegenstand mehrerer Besprechungen gewesen. Da es sich um betriebliche Daten handele, sei es an den Klägern, etwaige Einwendungen zu substantiieren.

Entscheidungsgründe

II.

20 Die Revision ist begründet. Sie führt zur Aufhebung des angefochtenen Urteils und zur Zurückverweisung der Sache zur anderweitigen Verhandlung und Entscheidung (§ 126 Abs. 3 Satz 1 Nr. 2 der Finanzgerichtsordnung --FGO--).

21 Das FG hat im Ergebnis zutreffend erkannt, dass wegen vorhandener Mängel in der Buchführung des Klägers dem Grunde nach eine Schätzungsbefugnis besteht (dazu unten 1.). Der Zeitreihenvergleich weist allerdings im Vergleich zu anderen Verprobungs- und Schätzungsmethoden besondere Problembereiche auf (dazu unten 2.), die seine Anwendbarkeit sowohl dem Grunde nach als auch hinsichtlich der Übernahme des sich aus einem Zeitreihenvergleich ergebenden "Mehrergebnisses" als Betrag der Hinzuschätzung der Höhe nach einschränken (unten 3.). Diese Einschränkungen hat das FG im Streitfall nicht in vollem Umfang beachtet (unten 4.), so dass die Sache an die Vorinstanz zurückgehen muss (unten 5.).

22 1. Das FA --und über § 96 Abs. 1 Satz 1 Halbsatz 2 FGO auch das FG-- war gemäß § 162 Abs. 2 Satz 2 Alternative 2 AO zur Schätzung befugt, weil die Buchführung des Klägers der Besteuerung nicht nach § 158 AO zugrunde gelegt werden kann.

23 a) Das FG hat die folgenden formellen Buchführungsmängel festgestellt, deren Vorliegen mittlerweile zudem zwischen den Beteiligten unstreitig ist, die allerdings von unterschiedlichem Gewicht sind:

24 aa) Die Unvollständigkeit der Tagesendsummenbons für die Zeit vom 1. Januar bis zum 16. März 2001 ist ein formeller Mangel, der --bezogen auf diesen Zeitraum-- grundsätzlich als gravierend anzusehen ist, da die vollständige Erfassung der Bareinnahmen nicht gesichert ist. Zwar konnten den Klägern konkrete Einnahmenverkürzungen nicht nachgewiesen werden. Sie haben ihre Behauptung, es handele sich um --für das Buchführungsergebnis unbeachtliche-- Leerbons, aber ebenfalls nicht belegen können.

25 bb) Die fehlende Datierung der Tagesendsummenbons des Jahres 2003 ist unter den Umständen des Streitfalls hingegen nur als eher geringfügiger formeller Mangel anzusehen. Die Bons sind lückenlos nummeriert, so dass sich daraus zumindest ihre zeitliche Reihenfolge zwingend ergibt. Der Kläger hat in seinen Aufzeichnungen die aufeinander folgenden Bons den aufeinander folgenden Öffnungstagen des Betriebs im Jahr 2003 zugeordnet. Zweifel an der Richtigkeit dieser Zuordnung hat auch das FA nicht vorgebracht.

26 cc) Auch das Fehlen der Programmierprotokolle der Registrierkasse stellt einen formellen Mangel dar. Anweisungen zur Kassenprogrammierung sowie insbesondere die Programmierprotokolle, die nachträgliche Änderungen dokumentieren, sind nach § 147 Abs. 1 Nr. 1 AO als "sonstige Organisationsunterlagen" aufbewahrungspflichtig. Dies hat die Finanzverwaltung schon lange vor den Streitjahren vertreten (z.B. Schreiben des Bundesministeriums der Finanzen --BMF-- vom 7. November 1995, BStBl I 1995, 738, Tz. VI.c, sowie Tz. 6 der diesem BMF-Schreiben

beigefügten Grundsätze ordnungsmäßiger DV-gestützter Buchführungssysteme; BMF-Schreiben vom 9. Januar 1996, BStBl I 1996, 34; zeitlich nach den Streitjahren auch BMF-Schreiben vom 26. November 2010, BStBl I 2010, 1342, und vom 14. November 2014, BStBl I 2014, 1450, Tz 111). Der erkennende Senat schließt sich dieser Auffassung an (ebenso Drüen in Tipke/Kruse, Abgabenordnung, Finanzgerichtsordnung, § 147 AO Rz 9, 26).

27 Die höchstrichterliche Rechtsprechung hatte bisher allerdings keine Gelegenheit, sich zu dem Gewicht dieses Mangels zu äußern. Der erkennende Senat vertritt hierzu die Auffassung, dass das Fehlen einer lückenlosen Dokumentation zur Kassenprogrammierung in seinen Auswirkungen auf die Beurteilung der formellen Ordnungsmäßigkeit der Buchführung und der Eröffnung der Schätzungsbefugnis dem Fehlen von Tagesendsummenbons bei einer Registrierkasse bzw. dem Fehlen täglicher Protokolle über das Auszählen einer offenen Ladenkasse gleichsteht. In allen drei Fällen lässt der formelle Mangel zwar keinen sicheren Schluss auf die Verkürzung von Einnahmen zu. Gleichwohl gibt es systembedingt keine Gewähr mehr für die Vollständigkeit der Erfassung der Bareinnahmen, ohne dass eine nachträgliche Ergänzung der Dokumentation bzw. eine anderweitige Heilung des Mangels möglich wäre. Elektronische Kassensysteme sind durch Umprogrammierung in nahezu beliebiger Weise manipulierbar; von derartigen Manipulationsmöglichkeiten machen Teile der betrieblichen Praxis nach dem Erkenntnisstand des Senats durchaus Gebrauch (zu einem solchen Fall z.B. Beschluss des FG Rheinland-Pfalz vom 7. Januar 2015 5 V 2068/14; vgl. zum Ganzen auch Tz. 54 der Bemerkungen des Bundesrechnungshofs 2003 zur Haushalts- und Wirtschaftsführung, BTDrucks 15/2020, 197 f.). Es ist daher von erheblicher Bedeutung, dass ein Betriebsprüfer --und ggf. auch ein FG-- sich davon überzeugen kann, wie die Kasse im Zeitpunkt ihrer Auslieferung und Inbetriebnahme programmiert war, sowie ob bzw. in welchem Umfang nach der Inbetriebnahme der Kasse spätere Programmeingriffe vorgenommen worden sind. Für den Steuerpflichtigen überschreitet der mit der Dokumentation verbundene Aufwand die Grenze des Zumutbaren nicht. Beim Erwerb der Kasse kann er vom Verkäufer die Übergabe von Bedienungsanleitungen und Programmdokumentationen verlangen. Die Dokumentation späterer Umprogrammierungen verursacht jedenfalls einen geringeren Aufwand als die Umprogrammierung selbst.

28 Das Gewicht dieses Mangels tritt allerdings zurück, wenn der Steuerpflichtige für den konkreten Einzelfall darlegt, dass die von ihm verwendete elektronische Kasse trotz ihrer Programmierbarkeit ausnahmsweise keine Manipulationsmöglichkeiten eröffnet.

29 dd) Keine sichere Grundlage bietet das angefochtene Urteil hingegen für die Beurteilung der Frage, welches Gewicht dem Fehlen von Kassenberichten für die Thekenkasse zukommt. Grundsätzlich ist die nur rechnerische Führung einer offenen Ladenkasse, die keine Kassensturzfähigkeit gewährleistet, als ein gravierender Mangel zu bewerten (vgl. Urteil des Bundesfinanzhofs --BFH-- vom 20. Juni 1985 IV R 41/82, BFH/NV 1985, 12), der schon für sich genommen auch ohne Nachweis einer konkreten *materiellen* Unrichtigkeit zu Hinzuschätzungen berechtigen würde. Allerdings hat das FG für das Jahr 2002 --in dem hinsichtlich der Thekenkasse dieselben Mängel bestanden wie für die Jahre 2001 und 2003-- auf Bl. 17 des Urteils ausdrücklich das Vorliegen von Buchführungsmängeln verneint und insoweit der Klage stattgegeben (auf Bl. 13 Abs. 2 des Urteils heißt es demgegenüber noch, die Kassenführung genüge "hinsichtlich der Streitjahre 2001 *bis* 2003" nicht den Anforderungen). Damit bleibt unklar, welche Bedeutung das FG den --vom Betriebsprüfer ohnehin nicht erwähnten-- Mängeln bei der Führung der Thekenkasse zugemessen hat. Revisionsrechtlich ist daher davon auszugehen, dass das FG diesem Gesichtspunkt keine entscheidende Bedeutung beigemessen hat.

30 ee) Ob die Speisekarte des Jahres 2001 im Streitfall gemäß § 147 Abs. 1 Nr. 5 AO aufbewahrungspflichtig war, kann anhand der Feststellungen des FG ebenfalls nicht beurteilt werden. Speisekarten sind nach der höchstrichterlichen Rechtsprechung nicht generell aufbewahrungspflichtig, sondern nur dann, wenn sie zum Verständnis und zur Überprüfung der gesetzlich vorgeschriebenen Aufzeichnungen im Einzelfall von Bedeutung sind (ausführlich hierzu BFH-Urteil vom 14. Dezember 2011 XI R 5/10, BFH/NV 2012, 1921, Rz 26 ff.). Ob diese Voraussetzung im Betrieb des Klägers erfüllt war, hat das FG nicht festgestellt. Letztlich kommt es auf diese Frage im Streitfall aber nicht an, da jedenfalls die übrigen festgestellten formellen Mängel für die Bejahung einer Schätzungsbefugnis ausreichen (dazu noch unten c).

31 b) Konkrete *materielle* Mängel der Kassenbuchführung, insbesondere der Einnahmenerfassung, hat das FG nicht festgestellt.

32 Zwar stellt die fehlende Inventur für sich genommen einen materiellen Mangel dar, da der in der Buchführung ausgewiesene --offenbar lediglich griffweise geschätzte-- Betrag des Warenendbestands nicht dem tatsächlichen Wert der vorhandenen Waren entspricht und dadurch zugleich der Jahresgewinn verfälscht wird (vgl. zum Gewicht

dieses Buchführungsmangels einerseits BFH-Urteil vom 14. Dezember 1966 VI 245/65, BFHE 87, 616, BStBl III 1967, 247; andererseits BFH-Urteil vom 26. August 1975 VIII R 109/70, BFHE 117, 224, BStBl II 1976, 210). Für die im vorliegenden Verfahren allein entscheidungserhebliche Frage der Ordnungsmäßigkeit der *Kassen*buchführung (Vollständigkeit der Einnahmenerfassung) ist dieser materielle Mangel indes ohne Bedeutung, da die Unrichtigkeit des Warenendbestands keine Wechselwirkung mit der Vollständigkeit der Erfassung der Bareinnahmen aufweist.

33 c) Aufgrund der festgestellten formellen Mängel entspricht die Buchführung des Klägers nicht den Vorschriften der §§ 140 bis 148 AO, so dass ihr nicht die Beweiskraftwirkung des § 158 AO zukommt. Dies würde nach dem Wortlaut des § 162 Abs. 2 Satz 2 Alternative 2 AO bereits zur Schätzung berechtigen.

34 Allerdings berechtigen formelle Buchführungsmängel nach ständiger höchstrichterlicher Rechtsprechung nur insoweit zur Schätzung, als sie Anlass geben, die sachliche Richtigkeit des Buchführungsergebnisses anzuzweifeln (BFH-Entscheidungen vom 17. November 1981 VIII R 174/77, BFHE 135, 11, BStBl II 1982, 430, unter 1.; vom 25. Januar 1990 IV B 140/88, BFH/NV 1990, 484, und in BFH/NV 2012, 1921, Rz 22, mit zahlreichen weiteren Nachweisen). Jedenfalls dann, wenn vorwiegend Bargeschäfte getätigt werden, können Mängel der Kassenführung aber der gesamten Buchführung die Ordnungsmäßigkeit nehmen (BFH-Urteil in BFH/NV 2012, 1921, Rz 34). Dies ist vorliegend der Fall. Das Fehlen der Programmierprotokolle der Registrierkasse ist --wie dargestellt-- als gewichtiger Mangel anzusehen, der den gesamten Streitzeitraum betrifft. Dass die Registrierkasse trotz ihrer Programmierbarkeit ausnahmsweise keine Manipulationsmöglichkeiten eröffnet, hat der Kläger nicht dargelegt. Für einen Teil des Jahres 2001 kommt noch das Fehlen eines Teils der Tagesendsummenbons als ebenfalls gewichtiger Mangel hinzu.

35 2. Zwar waren FA und FG danach im Streitfall zur Schätzung befugt. Die von ihnen gewählte Schätzungsmethode des Zeitreihenvergleichs (zur Beschreibung dieser Methode unten a) ist allerdings sowohl unter methodischen Aspekten (dazu unten b) als auch unter Rechtsschutzgesichtspunkten (unten c) nicht unproblematisch.

36 a) Der Zeitreihenvergleich stellt im Grundsatz eine mathematisch-statistische Verprobungsmethode dar; er ist dem Bereich der Strukturanalysen zuzurechnen (grundlegend zur Nutzung von Strukturanalysen in der Außenprüfung unter besonderer Berücksichtigung des Zeitreihenvergleichs Huber, Steuerliche Betriebsprüfung --StBp-- 2002, 199, 233, 258, 293; Wähnert, Neue Wirtschafts-Briefe 2012, 2774; Rau, Statistisch-mathematische Methoden der steuerlichen Betriebsprüfung und die Strukturanalyse als ergänzende Alternative, Köln 2012; speziell zum Zeitreihenvergleich Högemann, Die Information über Steuer und Wirtschaft 2000, 585; Wiggen, StBp 2008, 168; Vogelsang, BP-Handbuch, München 2008, I Rz 33 ff.; Brinkmann, Schätzungen im Steuerrecht, Berlin 2015, S. 278 ff.).

37 Prinzipiell sind zahlreiche Varianten eines Zeitreihenvergleichs denkbar. Die im vorliegenden Verfahren angewendete Methode stellt jedoch ausweislich der recht großen Zahl der hierzu ergangenen instanzgerichtlichen Entscheidungen die Standardherangehensweise der Finanzverwaltung dar (zu weiteren Varianten des Zeitreihenvergleichs vgl. Schumann/Wähnert, Steuerberatung 2012, 535). Sie ist dadurch gekennzeichnet, dass für jede Woche eines Kalenderjahres sowohl der wöchentliche Wareneinsatz als auch der Betrag der wöchentlichen Einnahmen ermittelt wird. Aus dem Vergleich dieser beiden Größen ergibt sich für jede Woche ein Rohgewinnaufschlagsatz. Die Finanzverwaltung sieht nun den höchsten Rohgewinnaufschlagsatz, der sich für einen beliebigen Zehn-Wochen-Zeitraum des Kalenderjahres ergibt, als maßgebend für das Gesamtjahr an. Da in der Buchhaltung des Steuerpflichtigen in aller Regel keine Daten vorhanden sind, aus denen sich die wöchentlichen Wareneinsätze entnehmen ließen, versucht die Finanzverwaltung, diese Größen im Wege der Schätzung aus einer Verteilung der aus den Eingangsrechnungen ersichtlichen Wareneinkäufe über den Zeitraum bis zum nächsten Einkauf gleichartiger Ware zu gewinnen. Ferner wird der Wareneinkauf um Personalbeköstigungen, Waren-Sachentnahmen und Warenbestandsveränderungen bereinigt.

38 b) Der Zeitreihenvergleich weist gegenüber anderen Verprobungs- und Schätzungsmethoden einige technisch bedingte Besonderheiten auf, die zumindest eine vorsichtige Interpretation seiner Ergebnisse gebieten.

39 aa) Auch bei einer formell und materiell ordnungsmäßigen Buchführung führt ein Zeitreihenvergleich denklogisch immer zu einem Mehrergebnis gegenüber der Buchführung, da der höchste Rohgewinnaufschlagsatz aller Zehn-Wochen-Perioden des Jahres auf den Wareneinsatz für das gesamte Jahr angewendet wird. Weil eine absolute Konstanz der wöchentlichen Rohgewinnaufschlagsätze in der Praxis auch bei Betrieben mit relativ geringer Lagerhaltung nicht vorkommt, muss der *höchste* Rohgewinnaufschlagsatz aller Zehn-Wochen-Perioden eines Jahres denknotwendig über dem *durchschnittlichen* Rohgewinnaufschlagsatz des Gesamtjahres liegen, selbst wenn dieser sich aus einer zutreffenden Buchführung ergibt (ebenso bereits FG Münster, Beschluss vom 19. August 2004

8 V 3055/04 G, EFG 2004, 1810, rechtskräftig). Andere Schätzungsmethoden (z.B. Aufschlagkalkulation, Geldverkehrsrechnung) weisen diese systembedingte Besonderheit nicht auf, da sie das Ergebnis einer zutreffenden Buchführung im Regelfall bestätigen --nicht aber widerlegen-- werden.

40 Daraus folgt, dass die vom FA in den Vordergrund seiner Argumentation gestellte Eignung des Zeitreihenvergleichs zur Feststellung von Doppelverkürzungen sowohl dem Grunde nach als auch zur Ermittlung ihres Umfangs der Höhe nach nicht frei von Zweifeln ist:

41 (1) Übersteigt der höchste Rohgewinnaufschlagsatz einer Zehn-Wochen-Periode den durchschnittlichen Rohgewinnaufschlagsatz des Jahres, steht allein mit diesem Befund grundsätzlich noch nicht mit der für eine Überzeugungsbildung nach § 96 Abs. 1 Satz 1 FGO erforderlichen Sicherheit fest, dass die rechnerische Differenz gerade auf dem Verschweigen von Erlösen beruht. Vielmehr lässt sich bei einem solchen Ergebnis --ohne das Vorliegen weiterer Anhaltspunkte-- zunächst einmal nicht ausschließen, dass es durch die dargestellte methodische Besonderheit des Zeitreihenvergleichs hervorgerufen sein könnte. Soweit das FA --und sinngemäß wohl auch das FG-- die Auffassung vertritt, ein sich aufgrund eines Zeitreihenvergleichs ergebendes Mehrergebnis lasse "kaum Raum für Zweifel daran, dass Erlöse nicht richtig verbucht seien" (ebenso FG Düsseldorf, Beschluss vom 15. Februar 2007 16 V 4691/06 A(E,U,F), EFG 2007, 814, rechtskräftig, unter II.2.d, und Urteil vom 20. März 2008 16 K 4689/06 E,U,F, EFG 2008, 1256, rechtskräftig, unter 3.d), kann der Senat dem daher nicht folgen.

42 (2) Auch auf der anderen Seite des Spektrums möglicher Manipulationen dürfte eine Verprobung mittels eines Zeitreihenvergleichs nicht unbesehen geeignet sein, jedenfalls den Einsatz einer technisch fortgeschrittenen Manipulationssoftware aufzudecken. Vielmehr ist eine derartige Software durchaus in der Lage, das Verhältnis zwischen den manipulierten Wareneinkäufen und den Erlösen so zu gestalten, dass die Rohgewinnaufschlagsätze auch im Wochenvergleich nur in einem Rahmen schwanken, der von Betriebsprüfern üblicherweise noch für plausibel gehalten wird.

43 (3) Ein empirischer Nachweis der Eignung dieser Verprobungsmethode zur sicheren Aufdeckung von Steuerverkürzungen --insbesondere der mit anderen Methoden nur schwer aufzudeckenden sog. "Doppelverkürzungen" sowohl der Wareneinkäufe als auch der Erlöse-- ist nach dem Kenntnisstand des Senats bisher nicht geführt worden. Insbesondere ist in den bisher von den Finanzgerichten veröffentlichten Entscheidungen allein aus einer sich mittels eines Zeitreihenvergleichs ergebenden Hinzuschätzung in keinem Fall ausdrücklich darauf geschlossen worden, dass der Steuerpflichtige Doppelverkürzungen begangen hat. Das FA hat in seiner Revisionserwiderung zwar --ohne Angabe von Nachweisen-- angeführt, der Zeitreihenvergleich werde von Wirtschaftsprüfern bereits langjährig zur betriebswirtschaftlichen Analyse eingesetzt. Welchem Ziel diese Analysen dienen sollen und wie die Belastbarkeit ihrer Ergebnisse von den beteiligten Fachkreisen eingeschätzt wird, hat es aber nicht vorgetragen.

44 bb) Zudem werden die Ergebnisse eines Zeitreihenvergleichs in erheblichem Umfang durch mathematische "Hebelwirkungen" beeinflusst.

45 (1) Dies verdeutlicht im Streitfall beispielhaft eine Betrachtung der Verhältnisse des Jahres 2001: Hier war für das FA und FG der durchschnittliche Rohgewinnaufschlagsatz der Kalenderwochen 23 bis 32 maßgebend. Dabei ergaben sich vor allem für die beiden am Ende dieser Zehn-Wochen-Periode liegenden Kalenderwochen 31 und 32 sehr hohe rechnerische Rohgewinnaufschlagsätze (412 % bzw. 360 %), die damit zugleich auch den Zehn-Wochen-Durchschnitt in die Höhe trieben. Zöge man stattdessen die Kalenderwochen 21 bis 30 heran, würde sich lediglich ein durchschnittlicher Rohgewinnaufschlagsatz von 209,60 % ergeben (statt der vom FA und FG herangezogenen 241,54 %). Hier hätte also bereits eine Verschiebung des Betrachtungszeitraums um nur zwei Wochen erhebliche Auswirkungen auf das Gesamtergebnis der Schätzung.

46 (2) Für das Streitjahr 2003 hat sich der Zeitraum von der 28. bis zur 37. Kalenderwoche vor allem wegen der zu Beginn dieses Zeitraums liegenden 28. Kalenderwoche als Zehn-Wochen-Zeitraum mit dem höchsten Rohgewinnaufschlagsatz des Jahres dargestellt (263,42 %). Für diese 28. Kalenderwoche ergibt sich wegen eines recht geringen Wareneinkaufs ein rechnerischer Rohgewinnaufschlagsatz von 340,68 %. Ordnet man demgegenüber den Wareneinkauf der Vorwoche (27. Kalenderwoche: 1.552,46 €) zu 50 % der 28. Kalenderwoche zu, ergäbe sich die folgende Berechnung:

- neuer Wareneinsatz der Kalenderwochen 28 bis 37/2003: 14.030,65 € + 50 % von 1.552,46 € = 14.806,88 €

- neuer Rohgewinnaufschlagsatz der Kalenderwochen 28 bis 37/2003: 244,37 %

- Brutto-Hinzuschätzung: 31.752,96 € (statt 45.728,46 €; d.h. 31 % weniger als bisher).

47 Dies zeigt, dass bereits eine Veränderung in der Zuordnung der Einkäufe nur einer einzigen Woche erhebliche Auswirkungen auf das Gesamtergebnis der Zehn-Wochen-Periode hat. Noch größere Auswirkungen würden sich an dieser Stelle zeigen, wenn der außergewöhnlich hohe Wareneinkauf der 26. Kalenderwoche (2.807,88 €) beispielsweise zu 1/3 der 28. Kalenderwoche zugeordnet würde.

48 (3) Nun ist dem Senat durchaus bewusst, dass der Zeitreihenvergleich gerade darauf basiert, das höchste --und nicht etwa das zweit- oder dritthöchste oder ein noch anderes-- Ergebnis aller Zehn-Wochen-Perioden eines Jahres als maßgebend für das Gesamtjahr anzusehen. Die aufgezeigte Hebelwirkung insbesondere hoher rechnerischer Wochen-Rohgewinnaufschlagsätze am Anfang oder Ende des herausgegriffenen Zehn-Wochen-Zeitraums macht aber deutlich, wie wesentlich die sorgfältige Ermittlung der Schätzungsgrundlagen --vor allem des Wareneinkaufs und seiner Verteilung zum Zwecke der Gewinnung des Wareneinsatzes-- gerade am Anfang und am Ende dieses Zeitraums für eine methodisch korrekte Durchführung eines Zeitreihenvergleichs ist (dazu noch unten 3.c; vgl. hierzu auch Beschluss des FG Münster vom 11. Februar 2000 9 V 5542/99 K,U,F, Deutsches Steuerrecht/Entscheidungsdienst 2000, 549, rechtskräftig). Dabei ist hinsichtlich des systematischen Verständnisses des Zeitreihenvergleichs darauf hinzuweisen, dass der --für diese Schätzungsmethode grundlegende-- Wareneinsatz eben nicht den eigenen Buchführungsdaten des Steuerpflichtigen entnommen wird, sondern diese Größe im Wege der mehr oder weniger ungenauen Schätzung aus einer Verteilung des Wareneinkaufs abgeleitet wird (ebenso bereits zutreffend FG Köln, Urteil vom 27. Januar 2009 6 K 3954/07, EFG 2009, 1092, rechtskräftig, unter II.3.). Anders wäre es nur, wenn der gesamte Warenbestand durch ein ausgefeiltes Warenwirtschaftssystem verwaltet würde, was in den meisten Betrieben der "Bargeldbranche" aber nicht der Fall ist.

49 c) Auch unter dem Gesichtspunkt des Anspruchs der Steuerpflichtigen auf einen effektiven Rechtsschutz durch die Gerichte (Art. 19 Abs. 4, Art. 20 Abs. 3 GG) ist der Zeitreihenvergleich nicht frei von methodenspezifischen Bedenken. Schätzungsgrundlagen müssen von der Finanzbehörde so dargelegt werden, dass ihre Nachprüfung --insbesondere eine Schlüssigkeitsprüfung des zahlenmäßigen Ergebnisses der Schätzung-- möglich ist (BFH-Urteil in BFH/NV 2012, 1921, Rz 24). Dabei müssen sowohl die Kalkulationsgrundlage --und damit auch die spezifischen Daten, auf denen der Zeitreihenvergleich basiert-- als auch die Ergebnisse der Kalkulation sowie die Ermittlungen, die zu diesen Ergebnissen geführt haben, offengelegt werden (BFH-Urteile vom 31. Juli 1974 I R 216/72, BFHE 113, 400, BStBl II 1975, 96, unter 2.b, und vom 17. November 1981 VIII R 174/77, BFHE 135, 11, BStBl II 1982, 430, unter 4.c).

50 aa) Für die Steuerpflichtigen, ihre Berater und auch die Finanzgerichte sind die Ergebnisse mathematisch-statistischer Methoden, zu denen auch der Zeitreihenvergleich gehört, wegen der dabei anfallenden umfangreichen Datenmengen nur beschränkt nachprüfbar. So umfasst im Streitfall allein die vom Betriebsprüfer erstellte --und für die Richtigkeit der Ergebnisse des Zeitreihenvergleichs grundlegende-- Aufgliederung der Wareneinkäufe bereits für den relativ kleinen Betrieb des Klägers und nur für das Streitjahr 2001 eine Excel-Tabelle mit ca. 1 100 Zeilen zu je 10 Spalten, d.h. insgesamt ca. 11 000 Eintragungen. Die Nachvollziehbarkeit dieser Datenaufbereitung des Prüfers wird noch dadurch erschwert --bzw. in Teilbereichen sogar ausgeschlossen--, dass die Tabelle weder chronologisch noch nach den einzelnen Lieferanten geordnet ist. Für das Gericht ist daher weder nachprüfbar, ob der Prüfer den Wareneinkauf zutreffend auf die einzelnen Kalenderwochen verteilt hat, noch ist --hier mit Ausnahme allenfalls des Biereinkaufs-- erkennbar, ob sich bestimmte Einkaufsmuster im Betrieb des Klägers regelmäßig wiederholen oder nicht.

51 Hinzu kommt, dass das FA im Streitfall noch nicht einmal das vollständige Zahlenwerk vorgelegt hat. Insbesondere ist aus den vorliegenden Unterlagen nicht ersichtlich, wie der Prüfer den sich aus den Eingangsrechnungen ergebenden Wareneinkauf auf die einzelnen Kalenderwochen verteilt hat, obwohl vor allem dieser Punkt zwischen den Beteiligten umstritten ist. Gerade die Verteilung des Wareneinkaufs durch den Prüfer --bzw. durch die eingesetzte Software-- wird im Regelfall aber den Schlüssel zum Verständnis und zur Einordnung der Einzelergebnisse des Zeitreihenvergleichs darstellen. Die Kenntnis der bei diesem Schätzungsschritt vom Prüfer notwendigerweise vorgenommenen Wertungen --Datenmaterial zur wochenweisen Verteilung des Wareneinsatzes existiert im Betrieb in aller Regel nicht, zumal es keine gesetzliche Verpflichtung zur Vornahme entsprechender Aufzeichnungen gibt-- ist für den Steuerpflichtigen von erheblicher Bedeutung, um Fehler oder Unsicherheiten in der vom Prüfer vorgenommenen Verteilung aufzeigen zu können. Derartige Fehler können --insbesondere wenn sie

dem Prüfer am Anfang oder Ende der von ihm herausgegriffenen Zehn-Wochen-Periode unterlaufen-- aufgrund des aufgezeigten mathematischen Hebeleffekts das rechnerische Ergebnis des Zeitreihenvergleichs in erheblichem Umfang beeinflussen und verzerren. Auch für das Gericht ist der Einblick in den Verteilungsvorgang wesentlich, um Erkenntnisse über die Größenordnung der im konkreten Fall anzunehmenden Fehlermarge des Zeitreihenvergleichs zu gewinnen.

52 bb) Die Finanzbehörde versetzt sich durch die Anwendung solcher mathematisch-statistischer Methoden aufgrund des ihnen innewohnenden Datenüberschusses daher in eine gewisse technisch-rechnerische Überlegenheit gegenüber dem Steuerpflichtigen. Dieser steht nun --jedenfalls nach Auffassung der Verwaltung-- in der Pflicht, "Auffälligkeiten" in den Ergebnissen des Zahlenwerks zu erklären bzw. zu widerlegen, verfügt aber, ohne dass ihm dies rechtlich vorzuwerfen wäre, möglicherweise gar nicht über das umfangreiche Zahlenmaterial --oder auch über das statistisch-methodische Wissen--, das erforderlich wäre, um eine sachgerechte Analyse der Datenmengen vornehmen zu können.

53 cc) Die vorstehenden Ausführungen zeigen, dass in Fällen der Anwendung eines Zeitreihenvergleichs die Finanzämter häufig schon ihre Darlegungspflichten nicht erfüllen, indem sie den Steuerpflichtigen und Finanzgerichten nicht alle Daten zur Verfügung stellen, die für eine vollständige Überprüfung des Zahlenwerks erforderlich sind. Auf der anderen Seite verbleiben aber selbst bei vollständiger Aufdeckung des Zahlenwerks rechtsstaatliche Bedenken, da die dann gelieferten Datenmengen so groß sind, dass eine grundlegende Überprüfung durch die Gerichte kaum zu leisten sein dürfte.

54 3. Aus diesen Befunden --so sensibel sie unter den aufgezeigten Aspekten der folgerichtigen Methodik und des effektiven Rechtsschutzes sein mögen-- folgt für den Senat auf der Grundlage des derzeitigen Erkenntnisstandes allerdings nicht das Ergebnis, die Methode des Zeitreihenvergleichs grundsätzlich zu verwerfen. Vielmehr besteht auf Seiten der Finanzverwaltung ein durchaus nachvollziehbares Bedürfnis, moderne Prüfungsmethoden --zu denen in geeigneten Fällen auch mathematisch-statistische Methoden gehören können-- einzusetzen. Denn auch umgekehrt verschafft sich ein Teil der Steuerpflichtigen durch ausgefeilt geplante Doppelverkürzungen und/oder den Einsatz von Manipulationssoftware, die die Aufdeckungsmöglichkeiten herkömmlicher Prüfungsmethoden minimieren, technische Vorteile gegenüber der Finanzverwaltung (vgl. dazu bereits oben 1.a cc). Hierauf darf und muss die Außenprüfung im Interesse der Wahrung der Gesetzmäßigkeit und Gleichmäßigkeit der Besteuerung (Art. 3 Abs. 1, Art. 20 Abs. 3 GG) reagieren, und zwar auch im Wege der Entwicklung und Anwendung neuartiger Prüfungsmethoden.

55 Die aufgezeigten Problembereiche lassen allerdings erkennen, dass der Zeitreihenvergleich für bestimmte Betriebstypen oder bestimmte betriebliche Situationen schon dem Grunde nach keine geeignete Schätzungsmethode darstellt (dazu unten a). Sofern der Zeitreihenvergleich dem Grunde nach als geeignete Methode anzusehen ist, sind seine Ergebnisse --in Abhängigkeit vom Grad der Fehlerhaftigkeit der Buchführung des Steuerpflichtigen und vom Umfang der im jeweiligen Einzelfall in Bezug auf die vollständige Erfassung der Erlöse vorliegenden sonstigen Erkenntnisse-- ggf. nur eingeschränkt für die Höhe der Hinzuschätzung zu übernehmen (unten b). In jedem Fall ist wegen der erheblichen Hebelwirkung methodischer Fehler, die bei der Durchführung des Zeitreihenvergleichs unterlaufen, auf eine besonders sorgfältige Ermittlung der Tatsachengrundlagen zu achten (unten c).

56 a) Bei bestimmten Betriebstypen oder in bestimmten betrieblichen Situationen scheidet der Zeitreihenvergleich schon dem Grunde nach als geeignete Schätzungsmethode aus. So basiert er entscheidend auf der Grundannahme, dass im Betrieb des Steuerpflichtigen das Verhältnis zwischen dem Wareneinsatz und den Erlösen im betrachteten Zeitraum --der allerdings nicht notwendig ein volles Kalenderjahr umfassen muss (z.B. bei Saisonbetrieben)-- weitgehend konstant ist. Fehlt es an dieser weitgehenden Konstanz, haben die Ergebnisse eines Zeitreihenvergleichs regelmäßig keine hinreichende Aussagekraft.

57 Auch darf es im maßgebenden Zeitraum nicht zu solchen Änderungen in der Betriebsstruktur gekommen sein, die --nicht anderweitig behebbare-- wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringen.

58 Diese Einschränkungen sind --abstrakt gesehen-- weitestgehend unstreitig und werden, soweit für den Senat ersichtlich, von der Finanzverwaltung im Allgemeinen auch schon bisher bei der Anwendung dieser Verprobungsmethode beachtet.

59 b) Sofern der Zeitreihenvergleich dem Grunde nach eine geeignete Verprobungsmethode darstellt, kann er

gleichwohl gegenüber anderen Methoden nachrangig sein bzw. können seine Ergebnisse nur unter Beachtung der nachfolgend (unter aa bis cc) dargestellten Abstufungen der Schätzung zugrunde gelegt werden.

60 Rechtliche Grundlage dieser Einschränkungen ist die Vorschrift des § 5 AO in Verbindung mit dem Grundsatz der Verhältnismäßigkeit, da die Wahlfreiheit des FA bei der Auswahl zwischen mehreren in Betracht kommenden Schätzungsmethoden nach den für die Ausübung pflichtgemäßen Ermessens geltenden Grundsätzen eingeschränkt ist und dabei auch Verhältnismäßigkeitserwägungen zu beachten sind. Jede Schätzung hat zum Ziel, Besteuerungsgrundlagen mit Hilfe von Wahrscheinlichkeitsüberlegungen zu ermitteln, wenn eine sichere Tatsachenfeststellung trotz des Bemühens um Aufklärung nicht möglich ist (BFH-Urteil vom 2. Februar 1982 VIII R 65/80, BFHE 135, 158, BStBl II 1982, 409, unter 1.c). Die Auswahl zwischen verschiedenen Schätzungsmethoden steht grundsätzlich im pflichtgemäßen Ermessen des FA bzw. FG (vgl. BFH-Beschluss vom 3. September 1998 XI B 209/95, BFH/NV 1999, 290, unter II.2.b). Ermessensleitend ist dabei das Ziel, die Besteuerungsgrundlagen durch Wahrscheinlichkeitsüberlegungen so zu bestimmen, dass sie der Wirklichkeit möglichst nahe kommen (BFH-Urteil vom 18. Dezember 1984 VIII R 195/82, BFHE 142, 558, BStBl II 1986, 226, unter 2.). Kommt eine bestimmte Schätzungsmethode diesem Ziel voraussichtlich näher als eine andere, ist die erstgenannte unter Ermessensgesichtspunkten vorzugswürdig.

61 Darin liegt keine Abweichung von der --vom FA angeführten-- Rechtsprechung, wonach der Steuerpflichtige grundsätzlich keinen Anspruch auf die Anwendung einer bestimmten Schätzungsmethode hat (vgl. hierzu Senatsbeschlüsse vom 1. März 2005 X B 158/04, BFH/NV 2005, 1014, unter 2.a, und vom 27. Januar 2009 X B 28/08, BFH/NV 2009, 717, unter 3.b). Denn dies lässt die Geltung der Grundsätze für die Ausübung pflichtgemäßen Ermessens unberührt. Im Übrigen betrafen diese Entscheidungen Fälle, in denen der Steuerpflichtige begehrte, das Ergebnis einer ordnungsgemäß angewendeten Schätzungsmethode durch Anwendung einer anderen, jedoch nicht vorrangigen oder besser geeigneten Methode zu überprüfen (vgl. BFH-Beschluss in BFH/NV 1999, 290, unter II.2.b, und in BFH/NV 2009, 717, unter 3.b: keine Überprüfung einer Aufschlagkalkulation durch eine Geldverkehrs- bzw. Vermögenszuwachsrechnung erforderlich; BFH-Beschluss in BFH/NV 2005, 1014, unter 2.a: keine Überprüfung einer Geldverkehrsrechnung durch eine Nachkalkulation erforderlich). Darum geht es vorliegend indes nicht.

62 aa) Bei einer Buchführung, die formell ordnungsgemäß ist oder --wie in dem Fall, der dem Urteil des FG Köln in EFG 2009, 1092 zugrunde lag-- nur geringfügige formelle Mängel aufweist, kann grundsätzlich nicht *allein* aufgrund der Ergebnisse eines Zeitreihenvergleichs der Nachweis der materiellen Unrichtigkeit geführt werden. Diese Einschränkung ist notwendige Folge des Befunds, dass ein Zeitreihenvergleich auch bei einer formell und materiell ordnungsmäßigen Buchführung stets zu einem rechnerischen "Mehrergebnis" führen wird, ein solches rechnerisches Mehrergebnis allein also kein hinreichendes Indiz für eine unvollständige Erfassung von Einnahmen darstellen kann (dazu oben 2.b aa). Schon die bisherige höchstrichterliche Rechtsprechung ist von dem Grundsatz ausgegangen, dass die Richtigkeitsvermutung einer formell ordnungsmäßigen Buchführung nur entkräftet ist, wenn das FA nachweist, dass das Buchführungsergebnis sachlich schlechterdings nicht zutreffen kann; an die Methodik einer solchen Schätzung sind wesentlich strengere Anforderungen zu stellen als in Fällen, in denen wegen festgestellter Buchführungsmängel ohnehin eine Schätzung der Einnahmen durchgeführt werden muss (vgl. BFH-Urteil vom 22. August 1985 IV R 29-30/84, BFH/NV 1986, 719, unter 2. vor a).

63 bb) Ist die Buchführung formell nicht ordnungsgemäß, sind aber materielle Unrichtigkeiten der Einnahmenerfassung nicht konkret nachgewiesen --so ist der Streitfall nach den Feststellungen des FG gelagert--, ist das FA gemäß § 162 Abs. 2 Satz 2 AO zwar dem Grunde nach zur Vornahme von Hinzuschätzungen berechtigt, weil die Richtigkeitsvermutung des § 158 AO nicht gilt. Allein die Ergebnisse eines Zeitreihenvergleichs lassen aufgrund der dieser Verprobungsmethode innewohnenden methodenbedingten Unsicherheiten aber noch keinen *sicheren* Schluss auf das Vorliegen und den Umfang auch materieller Unrichtigkeiten der Buchführung zu.

64 In diesen Fällen sind andere Schätzungsmethoden, die auf betriebsinternen Daten aufbauen oder in anderer Weise die individuellen Verhältnisse des jeweiligen Steuerpflichtigen berücksichtigen (z.B. Vermögenszuwachs- oder Geldverkehrsrechnung, Aufschlagkalkulation) grundsätzlich vorrangig heranzuziehen. Nur wenn solche Schätzungsmethoden --etwa wegen fehlender Mitwirkung des Steuerpflichtigen oder eines zu hohen Grades an Komplexität seiner betrieblichen oder sonstigen finanziellen Verhältnisse-- nicht sinnvoll einsetzbar sind, können die Ergebnisse eines Zeitreihenvergleichs einen Anhaltspunkt für die Höhe der erforderlichen Hinzuschätzung bilden.

65 Diese Ergebnisse sind vom FA und FG aber --auch von Amts wegen-- stets auf ihre Plausibilität anhand der

besonderen betrieblichen Verhältnisse des Steuerpflichtigen zu überprüfen, soweit diese bekannt sind (ebenso Schuster in Hübschmann/Hepp/ Spitaler, § 158 AO Rz 17). Bei verbleibenden Zweifeln können Sicherheitsabschläge in einem Umfang geboten sein, der über eine bloße Abrundung des "Mehrergebnisses" hinausgeht, das sich rechnerisch bei Anwendung des höchsten Zehn-Wochen-Rohgewinnaufschlagsatzes darstellt.

66 cc) Steht hingegen bereits aus *anderen* Gründen fest, dass die Buchführung nicht nur formell, sondern auch materiell unrichtig ist (z.B. nicht gebuchte Wareneinkäufe, nachweislich unversteuerte Betriebseinnahmen, rechnerische Kassenfehlbeträge usw.), und übersteigt die nachgewiesene materielle Unrichtigkeit der Buchführung eine von den Umständen des Einzelfalls abhängige Bagatellschwelle, können die Ergebnisse eines --technisch korrekt durchgeführten-- Zeitreihenvergleichs auch für die Ermittlung der erforderlichen Hinzuschätzung der *Höhe* nach herangezogen werden, sofern sich im Einzelfall keine andere Schätzungsmethode aufdrängt, die tendenziell zu genaueren Ergebnissen führt und mit vertretbarem Aufwand einsetzbar ist.

67 c) Hinsichtlich seiner technischen Durchführung setzt der Zeitreihenvergleich als mathematisch-statistische Methode eine besonders sorgfältige Ermittlung der Tatsachengrundlagen (Ausgangsparameter) voraus, zumal schon nach § 162 Abs. 1 Satz 2 AO bei der Schätzung alle Umstände zu berücksichtigen sind, die für sie von Bedeutung sind. Jedenfalls die Übernahme der vorhandenen Buchhaltungsdaten durch den Prüfer muss frei von Fehlern sein. Zwar gehen auch bei dieser Schätzungsmethode Unsicherheiten, die auf unzureichenden Aufzeichnungen des Steuerpflichtigen beruhen, zu dessen Lasten (allgemein hierzu BFH-Urteil in BFHE 142, 558, BStBl II 1986, 226, unter 2.). Da solche Unsicherheiten aber aufgrund der dem Zeitreihenvergleich innewohnenden Hebelwirkung erheblich verstärkt auf das Schätzungsergebnis "durchschlagen", ist in derartigen Fällen auch dann, wenn die Unsicherheiten auf einer Verletzung der Mitwirkungspflichten des Steuerpflichtigen beruhen, jedenfalls eine Plausibilitätsprüfung der Ergebnisse des Zeitreihenvergleichs vorzunehmen, die sich nicht allein auf einen summarischen Vergleich mit den amtlichen Richtsätzen beschränken darf.

68 Ferner wird das Zahlenwerk des Prüfers im Regelfall die Umschlaghäufigkeit der einzelnen Waren bzw. Warengruppen erkennen lassen müssen. Dies ist erforderlich, um eine hinreichend tragfähige Einschätzung zu gewinnen, ob die Grundvoraussetzung für die sinnvolle Durchführung eines Zeitreihenvergleichs --eine zu vernachlässigende bzw. zumindest hinreichend konstante Lagerhaltung-- erfüllt ist. Auch muss der vorhandene Warenbestand in sachgerechter Weise, die vor allem die Besonderheiten des geprüften Betriebs berücksichtigt, in die Betrachtung einbezogen werden.

69 Sollten die vorstehenden Anforderungen im Einzelfall nicht beachtet werden können, ist zumindest eine Vergleichsrechnung (Sensitivitätsanalyse) anzustellen. Diese muss verdeutlichen, welche Auswirkungen die nicht behebbaren Unsicherheiten bei einzelnen Parametern --vor allem solche Unsicherheiten, die darauf beruhen, dass der Zeitreihenvergleich auf einem Vergleich der wöchentlichen Wareneinsätze beruht, zu deren exakter Ermittlung der Steuerpflichtige aber von Gesetzes wegen nicht verpflichtet ist-- auf die Ergebnisse des Zeitreihenvergleichs haben können. Eine solche Sensitivitätsanalyse gehört dann, wenn sie durch vorhandene Unsicherheiten geboten ist, bereits zu den *formellen* Anforderungen, die an die technisch korrekte Durchführung des Zeitreihenvergleichs zu stellen sind. Sie ist daher vom FA von Amts wegen durchzuführen und vorzulegen, damit der Steuerpflichtige, sein Berater, das FG, aber auch das FA selbst den Umfang der im Einzelfall möglichen Fehlermarge einschätzen können. Allerdings ist das FA --in Abhängigkeit vom Gewicht der formellen und der schon ohne den Zeitreihenvergleich zutage tretenden materiellen Mängel der Buchführung des Steuerpflichtigen-- nicht gehalten, sein Schätzungsergebnis stets an der zugunsten des Steuerpflichtigen äußersten Grenze der Fehlermarge zu orientieren.

70 4. Im Streitfall hat das FG bei seiner Entscheidung nicht alle rechtlichen Anforderungen beachtet, die an die Anwendung und Durchführung eines Zeitreihenvergleichs zu stellen sind.

71 a) Das FG hat sich nicht mit der Frage auseinandergesetzt, ob die (erneute) Eröffnung des Betriebs des Klägers im Jahr 2003 als Änderung in der Betriebsstruktur anzusehen ist, die wesentliche Unsicherheiten bei der Aufstellung und Interpretation des Zahlenwerks mit sich bringt (vgl. die Ausführungen oben 3.a).

72 aa) Insbesondere der Umstand, dass im Zeitpunkt der Betriebseröffnung ausweislich der vom Betriebsprüfer zugrunde gelegten Zahlen kein Warenbestand vorhanden war, hätte Anlass geben müssen, die Eignung des Zeitreihenvergleichs im konkreten Fall --zumindest aber die vom FA vorgenommene gleichmäßige Verteilung des Aufbaus des zum ersten folgenden Bilanzstichtag vorgefundenen Warenendbestands auf sämtliche Kalenderwochen zwischen der Betriebseröffnung und dem Jahresende-- einer kritischen Würdigung zu unterziehen. Dass die Art und Weise der Verteilung des Warenbestandsaufbaus von erheblicher Bedeutung für das Ergebnis des Zeitreihenvergleichs ist, folgt schon aus dem Umstand, dass der (geschätzte) Warenendbestand von 10.800 € sich

auf immerhin 13 % des gesamten Wareneinsatzes von 84.858,93 € beläuft. Ein Unsicherheitsfaktor im Umfang von 13 % des Wareneinsatzes kann aufgrund der erheblichen Hebelwirkung dieses Berechnungsfaktors aber zu Schwankungsbreiten der einzelnen Wochenwerte führen, die die Verwertbarkeit der Ergebnisse des Zeitreihenvergleichs insgesamt in Frage stellen.

73 Zwar ist die unterbliebene Inventur als eine Verletzung von Mitwirkungspflichten anzusehen, aus der der Kläger grundsätzlich keine Vorteile ziehen darf. Hinsichtlich der Verteilung des Warenbestandsaufbaus auf die einzelnen Wochen des Jahres gibt es jedoch keine besonderen Aufzeichnungs- oder sonstige Mitwirkungspflichten. FA und FG sind daher auch im Falle des pflichtwidrigen Unterlassens einer Inventur gehalten, den in der Bilanz ausgewiesenen Warenendbestand im Rahmen der Durchführung eines Zeitreihenvergleichs nach sachgerechten Kriterien auf die einzelnen Wochen des Jahres zu verteilen.

74 bb) Bei Betrachtung des --vom FG festgestellten und daher für den erkennenden Senat revisionsrechtlich in vollem Umfang verwertbaren-- Zahlenwerks des vom FA durchgeführten Zeitreihenvergleichs springt ins Auge, dass die vom Betriebsprüfer ermittelten Rohgewinnaufschlagsätze für die ersten zehn Wochen nach der Betriebseröffnung sehr niedrig sind. Dies deutet darauf hin, dass bereits in diesen Wochen ein erheblicher Teil des Warenbestands aufgebaut wurde. Eine solche Annahme erscheint auch in tatsächlicher Hinsicht als durchaus plausibel, zumal die vom FA und FG zugrunde gelegte abweichende Auffassung, der Warenbestand sei während des Jahres gleichmäßig bis zum Erreichen des in der Bilanz ausgewiesenen Endbestands angewachsen, für die ersten Wochen nach Betriebseröffnung zu Ergebnissen führen würde, die nicht realitätsgerecht erscheinen. So hätte der Kläger z.B. in der ersten Woche nach der Betriebseröffnung mit einem Warenbestand im Wert von lediglich 245 € wirtschaften und auskommen müssen, was --insbesondere angesichts der von der Kundschaft eines Gastronomiebetriebs dieser Art erwarteten Auswahl zwischen verschiedenen Alkoholika-- auszuschließen sein dürfte. Wird aber ein größerer Teil des Warenbestandsaufbaus als bisher den ersten Wochen nach der Betriebseröffnung zugeordnet, würden die --sehr niedrigen-- vom Prüfer für diesen Zeitraum errechneten Rohgewinnaufschlagsätze steigen. Korrespondierend dazu würden sich die Rohgewinnaufschlagsätze sämtlicher Folgewochen --darunter auch die Werte in den vom FA herangezogenen Kalenderwochen 28 bis 37-- entsprechend mindern, was einen unmittelbaren und erheblichen Einfluss auf das Schätzungsergebnis hätte.

75 cc) Werden nun beispielsweise 70 % des gesamten Warenbestandsaufbaus den ersten zehn Wochen nach der Betriebseröffnung zugeordnet und die verbleibenden 30 % des Bestandsaufbaus gleichmäßig auf die restlichen Wochen des Jahres 2003 verteilt --was zwar ebenfalls eine Sachverhaltsunterstellung und damit eine Schätzung darstellt, die aber deutlich realitätsnäher sein dürfte als die vom FA und FG vorgenommene Verteilung--, ergibt sich die folgende Neuberechnung:

76 (1) Der Betriebsprüfer hat für die ersten zehn Wochen nach der Betriebseröffnung (Kalenderwochen 10 bis 19/2003) einen Wareneinsatz von 19.536,01 € zugrunde gelegt. Im Wege des Vergleichs mit den in diesen Wochen erzielten Einnahmen (45.019,02 €) hat er einen Zehn-Wochen-Durchschnitts-Rohgewinnaufschlagsatz von 130,44 % ermittelt (45.019,02 € ÷ 19.536,01 € ./. 100 %). Bei Ermittlung des genannten Wareneinsatzes hat der Prüfer einen Warenbestandsaufbau im Umfang von 10 x 245,45 € (insgesamt 2.454,50 €) mindernd berücksichtigt. Ohne Minderung um den verteilten Warenendbestand hätte sich der Wareneinsatz der Kalenderwochen 10 bis 19/2003 daher auf 21.990,51 € belaufen.

77 (2) Für den Zehn-Wochen-Zeitraum mit dem höchsten Rohgewinnaufschlagsatz (Kalenderwochen 28 bis 37/2003) hat der Prüfer einen Wareneinsatz von 14.030,65 € sowie Einnahmen von 50.990,05 € ermittelt. Daraus ergab sich der für das Gesamtjahr zugrunde gelegte höchste Rohgewinnaufschlagsatz von 263,42 %. Ohne Minderung um den gleichmäßig auf alle Wochen des Jahres verteilten Warenendbestand (2.454,50 €) hätte der Wareneinsatz sich auf 16.485,15 € belaufen.

78 (3) Werden nun 70 % des gesamten Warenbestandsaufbaus des Jahres auf die ersten zehn Kalenderwochen verteilt, ist insoweit ein Betrag von 7.560 € zu berücksichtigen (70 % von 10.800 €). Der Wareneinsatz dieser Periode würde sich auf 14.430,51 € belaufen (21.990,51 € unbereinigter Wareneinsatz ./. 7.560 € Warenbestandsaufbau). Angesichts der aufgezeichneten Einnahmen von 45.019,02 € ergibt sich ein Zehn-Wochen-Rohgewinnaufschlagsatz von 211,97 %. Dieser Wert erscheint erheblich plausibler als der vom Prüfer ermittelte --im Jahresvergleich äußerst geringe-- Wert von 130,44 %, für den weder das FA noch das FG eine Erklärung angeführt haben.

79 (4) Die verbleibenden 30 % (3.240 €) des Warenbestandsaufbaus sind nun auf die Kalenderwochen 20 bis 53/2003 zu verteilen (34 Wochen, d.h. 95,29 € pro Woche). Der Wareneinsatz für die Kalenderwochen 28 bis 37/2003 beläuft sich danach auf 15.532,25 € (16.485,15 € unbereinigter Wareneinsatz ./. 10 x 95,29 € Warenbestandsaufbau). Bei

Einnahmen von 50.990,05 € ergibt sich ein Rohgewinnaufschlagsatz von 228,29 %. Dieser Wert stellt zwar immer noch den höchsten aller Zehn-Wochen-Perioden des Jahres 2003 dar, liegt aber deutlich unterhalb des vom FA und FG angenommenen Wertes von 263,42 %.

80 (5) Der Betrag des rechnerischen Brutto-"Mehrergebnisses" für das Jahr 2003 würde sich damit schon allein aufgrund dieser einzelnen Korrektur von 45.728,46 € auf 19.955,32 € --also um über 56 %-- mindern.

81 b) Das FG hat ferner offensichtliche Fehler im Zahlenwerk des Betriebsprüfers unbeanstandet gelassen, obwohl die Durchführung des Zeitreihenvergleichs zwingend eine fehlerfreie Übernahme der vorhandenen Buchhaltungsdaten erfordert (vgl. dazu oben 3.c).

82 aa) So hat der Prüfer in seiner zusammenfassenden Berechnung für das Jahr 2001 vom Kläger erklärte Einnahmen in Höhe von 657.095,38 DM angesetzt (dies entspricht den in der Gewinn- und Verlustrechnung ausgewiesenen Werten ohne die Konten "Erlöse" und "Gutscheine" sowie ohne Sachentnahmen). Im Rahmen des Zeitreihenvergleichs hat er hingegen nur einen Gesamt-Einnahmenbetrag von 656.231,89 DM auf die einzelnen Kalenderwochen des Jahres 2001 verteilt. Bei korrekter Durchführung des Zeitreihenvergleichs müssen diese beiden Werte aber zwingend identisch sein.

83 Dieselbe Auffälligkeit ergibt sich für das Jahr 2003: Der Prüfer geht in der zusammenfassenden Berechnung von erklärten Erlösen in Höhe von 220.905,86 € aus (was mit den in der Gewinn- und Verlustrechnung ausgewiesenen Werten --ohne Sachentnahmen-- übereinstimmt), setzt im Zeitreihenvergleich aber nur Erlöse von 220.447,93 € an.

84 Diese vom FA nicht erläuterten Differenzbeträge --die nicht mehr mit bloßen Rundungsdifferenzen zu erklären sind-- deuten darauf hin, dass dem Prüfer bei der Erfassung der vom Kläger erklärten Erlöse oder bei ihrer Verteilung auf die einzelnen Kalenderwochen Fehler unterlaufen sind. Diese Fehler sind aufklärungsbedürftig; der erkennende Senat ist mangels Vorlage des vollständigen Rechenwerks durch das FA aber nicht in der Lage, die Fehlerursache selbst näher zu ergründen. Zwar sind die Differenzbeträge nicht allzu hoch. Sollten sie aber gerade die Zehn-Wochen-Zeiträume mit den jeweils höchsten Rohgewinnaufschlagsätzen der beiden Streitjahre betreffen, könnten sie aufgrund der dem Zeitreihenvergleich innewohnenden rechnerischen Hebelwirkung gleichwohl erhebliche Auswirkungen auf das Gesamtergebnis der Schätzung haben.

85 bb) Für das Jahr 2001 geht der Prüfer im Rahmen seines Zeitreihenvergleichs von einem Wareneinkauf (vor Korrektur um den Eigen- und Personalverbrauch) von 225.525,55 DM aus. In der Gewinn- und Verlustrechnung des Klägers ist hierfür allerdings ein Betrag von 227.614,05 DM ausgewiesen. Wäre dieser höhere Betrag im Rahmen des Zeitreihenvergleichs angesetzt worden, hätte sich zugunsten der Kläger ein geringerer Rohgewinnaufschlagsatz ergeben.

86 Gleiches gilt für das Jahr 2003: Der Prüfer hat einen Wareneinkauf von 94.347,15 € angesetzt, während sich aus der Gewinn- und Verlustrechnung (ohne Bestandsveränderung) ein Wert von 95.658,93 € ergibt.

87 cc) Für das Jahr 2001 ist die Vorgehensweise des Prüfers zudem insoweit inkonsequent, als er zwar in seiner zusammenfassenden Berechnung die zwischen dem Jahresanfang und dem Jahresende eingetretene Veränderung des Warenbestands berücksichtigt hat, nicht aber im wochenweisen Zahlenwerk des Zeitreihenvergleichs.

88 Zwar ist die betragsmäßige Auswirkung dieses Fehlers angesichts der nur geringen im Laufe des Jahres 2001 eingetretenen Bestandsveränderung nicht allzu bedeutsam. Die in mehrfacher Hinsicht erkennbar fehlende innere Konsistenz des Zahlenwerks des Prüfers hätte dem FG aber Anlass geben müssen, dieses *insgesamt* genauer zu überprüfen.

89 c) Ferner hat das FG die Berücksichtigung der vorgelegten Geldverkehrsrechnungen im Wesentlichen mit der Erwägung abgelehnt, ihnen lägen die sich aus der Buchführung ergebenden Betriebseinnahmen zugrunde, die aber wegen der fehlenden Ordnungsmäßigkeit der Buchführung nicht angesetzt werden dürften.

90 Diese Argumentation ist nicht frei von Denkfehlern. Wäre sie richtig, dann verbliebe für die --von der ständigen höchstrichterlichen Rechtsprechung anerkannte-- Methode der Geldverkehrsrechnung praktisch kein Anwendungsbereich mehr, da diese typischerweise gerade dann zum Einsatz kommt, wenn zumindest ein Anfangsverdacht der unvollständigen Erklärung der Erlöse besteht. Die Geldverkehrsrechnung dient in solchen Fällen gerade dem Nachweis, dass der im Prüfungszeitraum zu beobachtende Geldverkehr mit den erklärten (geringen) betrieblichen Erlösen nicht in Einklang zu bringen ist.

91 Wenn das FG aufgrund des Zeitreihenvergleichs davon überzeugt war, dass der Kläger tatsächlich höhere Betriebseinnahmen erzielt hatte als er in den Steuererklärungen und den von ihm vorgelegten Geldverkehrsrechnungen angegeben hatte, wären diese höheren Einnahmen --zur Vermeidung von Widersprüchen bei der Anwendung verschiedener Schätzungsmethoden-- auch in den Geldverkehrsrechnungen anzusetzen gewesen. Damit wären die frei verfügbaren Beträge aber noch deutlich höher ausgefallen als bereits in den von den Klägern vorgelegten Geldverkehrsrechnungen, was sich bei Zugrundelegung dieser Schätzungsmethode zugunsten der Kläger ausgewirkt hätte. Der Umstand, dass das FG vom Vorhandensein zusätzlicher Betriebseinnahmen überzeugt war, reicht daher allein noch nicht zur vollständigen Verwerfung der Geldverkehrsrechnung aus. Vielmehr hätte dieser Umstand es nahegelegt, die vorhandenen Geldverkehrsrechnungen anhand der eigenen Erkenntnisse des FG zur Höhe der Einnahmen des Klägers zu überarbeiten.

92 Hinzu kommt, dass in der Konstellation des Streitfalls --formelle Ordnungsmängel von einigem Gewicht, aber kein anderweitiger Nachweis konkreter materieller Mängel der Einnahmenerfassung-- Schätzungsmethoden wie die Geldverkehrsrechnung grundsätzlich vorrangig vor einem Zeitreihenvergleich heranzuziehen sind (dazu ausführlich oben 3.b bb). Dies gilt im Streitfall umso mehr, als sowohl die Kläger als auch das FA jeweils Geldverkehrsrechnungen samt dazugehöriger Unterlagen vorgelegt haben, die dem FG möglicherweise als Ausgangspunkt und Erkenntnismaterial für eine eigene Schätzung hätten dienen können.

93 5. Für das weitere Verfahren weist der Senat --ohne die Bindungswirkung des § 126 Abs. 5 FGO-- auf die folgenden Punkte hin:

94 a) Es bietet sich an, ergänzende Feststellungen dazu zu treffen, ob der Betrieb des Klägers von seiner Struktur her für die Durchführung eines aussagekräftigen Zeitreihenvergleichs geeignet ist. Insoweit ist aus den vom FA vorgelegten Unterlagen bisher lediglich erkennbar, dass der wichtigste Teil des Wareneinkaufs --der Getränkebezug von der Brauerei, an die der Kläger vertraglich gebunden war-- wöchentlich vorgenommen wurde, was auf die erforderliche Geringfügigkeit bzw. Konstanz der Lagerhaltung bei dieser Warengruppe hindeutet. Zum Einkaufsverhalten des Klägers in Bezug auf die anderen Warengruppen hat das FA bisher jedoch keine brauchbaren Unterlagen vorgelegt (vgl. dazu auch oben 3.c); entsprechend hat das FG hierzu keine Feststellungen treffen können.

95 b) Das FG kann prüfen, ob im Streitfall eine Geldverkehrsrechnung durchführbar erscheint.

96 aa) Hierfür könnte sprechen, dass sowohl die Kläger für sämtliche Streitjahre (sie haben zur Untermauerung ihres Vorbringens zudem zahlreiche Unterlagen vorgelegt) als auch das FA (wohl beschränkt auf das damalige weitere Streitjahr 2002) Geldverkehrsrechnungen durchgeführt und dem FG vorgelegt haben. Dies könnte darauf hindeuten, dass beide Beteiligten grundsätzlich von der Eignung dieser Verprobungs- und Schätzungsmethode für die sachgerechte Beurteilung des Streitfalls ausgehen.

97 bb) Dreh- und Angelpunkt einer Geldverkehrsrechnung wird die Finanzierung der im Jahr 2001 für 365.000 DM (zuzüglich der üblichen Nebenkosten) erworbenen selbstgenutzten Immobilie der Kläger sein. Die Kläger haben hierzu im Laufe des Verfahrens die folgenden Finanzierungselemente nachgewiesen:

- Bankdarlehen: 140.000 DM
- Auszahlung aus einem Lebensversicherungsvertrag: 79.794 DM
- Banküberweisung der im Ausland lebenden Schwester des Klägers (Gleichstellungsgeld im Rahmen einer Vermögensübertragung von der Mutter an die Schwester des Klägers): 90.813,26 DM

98 Damit ist ein Betrag von gut 310.000 DM abgedeckt. Zur Herkunft des Differenzbetrags haben die Kläger bisher keine Unterlagen vorgelegt, sondern sich auf die "Lebensversicherung" berufen. Diese ist aber bereits in der obigen Finanzierungsrechnung enthalten und kann daher nicht nochmals angesetzt werden. Dass es sich um eine Auszahlung aus einem weiteren, bisher nicht bekannten Lebensversicherungsvertrag handeln könnte, haben die Kläger weder selbst vorgetragen noch nachgewiesen.

99 Die Kläger haben nachgewiesen, dass der Betrag von 90.813,26 DM vom Bankkonto der im Ausland lebenden Schwester des Klägers überwiesen wurde. Das FG scheint indes weiterhin daran zu zweifeln, dass es sich tatsächlich um einen Mittelzufluss "von außen" handelt. In einem solchen Fall muss es den Steuerpflichtigen, der einen

solchen Mittelzufluss durch Vorlage von Bankbelegen nachgewiesen hat, vor einer ihm nachteiligen Entscheidung, die auf die verbleibenden Zweifel des FG gestützt wird, allerdings konkret auffordern, weitergehend auch z.B. nachzuweisen, wie derjenige, der dem Steuerpflichtigen den Geldbetrag überwiesen hat, zu diesen Mitteln gekommen ist (vgl. zu ähnlichen Sachverhalten Senatsbeschluss vom 12. Juni 2013 X B 191/12, BFH/NV 2013, 1622, Rz 16, m.w.N.). Eine solche Aufforderung ist bisher nicht ausdrücklich ergangen. Da der Mittelzufluss nach dem Vorbringen der Kläger im Zusammenhang mit einer Grundstücksübertragung von der Mutter auf die Schwester des Klägers gestanden haben soll und bei Grundstücksübertragungen erfahrungsgemäß auch nach längerer Zeit noch Unterlagen aufbewahrt werden, erscheint es nicht als ausgeschlossen, dass die Kläger auf eine im zweiten Rechtsgang noch konkret zu stellende Anfrage des FG Unterlagen aus dem Privatbereich der ausländischen Verwandten des Klägers werden vorlegen können, sofern das FG seine diesbezüglichen Zweifel aufrecht erhalten sollte.

100 cc) Da die Kläger sich und ihre Familie --wie die von ihnen aufgezeichneten Sachentnahmen zeigen-- in erheblichem Umfang aus der Gastwirtschaft des Klägers verpflegen, wären die von der Finanzverwaltung üblicherweise herangezogenen Beträge aus den Tabellen für die Lebenshaltungskosten durchschnittlicher Haushalte um diejenigen Positionen zu mindern, die im Fall der Kläger bereits durch Sachentnahmen abgedeckt sind.

101 dd) Sollte die Durchführung einer Geldverkehrsrechnung im Streitfall zum jetzigen Zeitpunkt an ihre Grenzen stoßen, weil die Vorgänge zu lange zurückliegen --die Kläger haben bereits im Jahr 2001 erhebliche Mittel für den Erwerb der von ihnen selbstgenutzten Wohnung verwendet; die Unterlagen, mit denen die Kläger das Vorhandensein von Eigenmitteln für den Immobilienerwerb zu belegen versucht haben, stammen sogar aus dem Jahr 2000--, hätte das FG zu erwägen, welche Folgen sich daraus ergäben. Nach den bereits dargelegten Grundsätzen (oben 3.b bb) könnte in einem solchen Fall die Heranziehung der Ergebnisse eines --technisch korrekt durchgeführten-- Zeitreihenvergleichs zulässig sein; von ihnen wären aber nennenswerte Abschläge vorzunehmen.

102 c) Während die allgemeinen methodischen Einwendungen der Kläger gegen den Zeitreihenvergleich im Ansatz berechtigt sind --und zu den vom Senat formulierten Anforderungen an die Wahl und die korrekte Durchführung dieser Verprobungs- und Schätzungsmethode geführt haben--, ist ihr Vorbringen zu den vermeintlichen Besonderheiten des konkreten Falles teils bereits rechtlich nicht erheblich und teils zu wenig substantiiert.

103 aa) Die Behauptung, der übliche Verderb eingekaufter Waren sei bisher nicht berücksichtigt worden, stellt schon im Ausgangspunkt keine geeignete Einwendung dar, weil auch verdorbene Ware im Wareneinkauf enthalten ist und --mangels Vornahme einer Korrektur-- in den Wareneinsatz eingeht. Dies gilt im Übrigen auch für die Ermittlung der amtlichen Richtsätze. Außerdem würde die Berücksichtigung des Warenverderbs zu deutlich *höheren* Rohgewinnaufschlagsätzen führen, weil der in der Buchhaltung ausgewiesene Wareneinkauf um den Wert der verdorbenen Waren zu mindern wäre und dadurch der Wareneinsatz sinken würde, was rechnerisch unmittelbar eine Erhöhung des Rohgewinnaufschlagsatzes bewirken würde.

104 bb) Die Kläger haben ihre Behauptung, in den letzten Wochen des Jahres 2003 sei ein besonders hoher Bierbestand aufgebaut worden, nicht anhand der --ihnen vorliegenden-- Buchhaltungszahlen zum Biereinkauf und den Biererlösen der letzten Wochen des Jahres substantiiert. Für das FG besteht daher derzeit kein Anlass, dieser Behauptung nachzugehen.

105 cc) Ferner haben die Kläger bisher nicht vermocht, konkrete Auswirkungen des X-Fests auf die Ergebnisse des Zeitreihenvergleichs darzulegen. Im Jahr 2001 lag das X-Fest nicht in dem vom FA herangezogenen Zehn-Wochen-Zeitraum mit dem höchsten Rohgewinnaufschlagsatz. Im Jahr 2003 lag es zwar in diesem Zehn-Wochen-Zeitraum; die entsprechende Woche weist auch die höchsten Erlöse des gesamten Jahres auf. Da das X-Fest aber weder am Anfang noch am Ende des Zehn-Wochen-Zeitraums lag, dürften sich wochenweise Verschiebungen zwischen dem besonderen Wareneinkauf für Zwecke des X-Fests und den zusätzlichen Erlösen ausgeglichen haben. Jedenfalls haben die Kläger über den bloßen Verweis auf das Stattfinden des X-Fests hinaus insoweit keine substantiierten Einwendungen gegen den Zeitreihenvergleich vorgebracht.

106 dd) Ebenso wenig substantiiert ist die Einwendung, im Streitjahr 2003 sei ein gerade kostendeckendes Mittagsgericht angeboten worden. Die Kläger haben weder eine Speisekarte vorgelegt, aus der sich ein besonders preiswertes Mittagsgericht ergäbe, noch haben sie dem FG eine Kalkulation zur Verfügung gestellt, die ihre Behauptung stützen würde, ein Gericht sei gerade kostendeckend gewesen. Ohnehin ist unklar, was die Kläger mit diesem Einwand bezwecken wollen, da der Zeitreihenvergleich --so er denn technisch korrekt durchgeführt wird-- hinsichtlich der Erlöse die eigenen Daten aus der Buchhaltung des Steuerpflichtigen zugrunde legt, sich also eine

Betriebsführung mit tatsächlich sehr geringen Rohgewinnaufschlagsätzen auch in den Ergebnissen des Zeitreihenvergleichs widerspiegeln würde.

107 ee) Preisschwankungen beim Ein- und Verkauf sowie saisonale Schwankungen in der Zusammensetzung der Ein- und Verkäufe können zwar erhebliche verzerrende Wirkungen auf die Ergebnisse eines Zeitreihenvergleichs haben. Über ihren --zutreffenden-- Hinweis auf derartige allgemeine Wirkmechanismen hinaus haben die Kläger aber nichts dazu vorgetragen, ob bzw. in welchem Umfang im konkreten Fall derartige Schwankungen eingetreten und daher zu berücksichtigen sind.

108 d) Hinsichtlich der Aussetzungszinsen, die das FG zusätzlich zum Abzug als Betriebsausgaben zugelassen hat, hat es im Tatbestand seiner Entscheidung ausgeführt, die Aussetzungszinsen seien für "Umsatzsteuernachzahlungen für 2001 und 2003" (d.h. die Streitjahre) entstanden. Nicht festgestellt ist, wann der Zinslauf begonnen hat. Nach Lage der Dinge dürfte der Zinslauf erst mit dem Einlegen eines Rechtsbehelfs --also nach Erlass der angefochtenen Änderungsbescheide im Jahr 2008-- begonnen haben (vgl. § 237 Abs. 2 AO). Das FG wird daher im zweiten Rechtsgang noch Feststellungen dazu treffen, ob es überhaupt denkbar ist, dass die Aussetzungszinsen wirtschaftlich bereits im Streitjahr 2003 verursacht sein können (zur wirtschaftlichen Verursachung von Steuerzinsen erst ab Beginn des Zinslaufs siehe auch BFH-Urteil vom 22. Dezember 2010 I R 110/09, BFHE 232, 415, BStBl II 2014, 119, Rz 29).

109 e) Sollte es im zweiten Rechtsgang zu einem (Teil-)Erfolg der Klage kommen, wäre die Gewerbesteuerrückstellung gegenläufig anzupassen. Dies hat das FG in seinem für das Streitjahr 2003 teilweise klagestattgebenden Urteil im ersten Rechtsgang übersehen.

110 6. Die Übertragung der Kostenentscheidung auf das FG beruht auf § 143 Abs. 2 FGO.

Quelle: www.bundesfinanzhof.de