Petra Ritzer-Angerer*

Was bedeuten die Vertrauensguteigenschaften der Jahresabschlussprüfung für die Regulierung der Wirtschaftsprüferhaftung?

<https://doi.org/10.1515/zfwp-2020-2031>

Abstract: The annual audit is the auditors' main activity. It includes risk analysis, planning and execution. Credence goods are special goods or services whose suppliers have superior information about the quality, e. g. therapies or repairs. The customer cannot evaluate whether the service was appropriate. As the audit is sophisticated and a professional act, it is reasonable to assume that the auditor is better informed. Hence the audit can be considered a credence good. This article discusses two questions: Does the audit comply with the characteristics of credence goods and can regulation – especially of the auditor's liability – mitigate the audit market problems caused by this?

JEL-Klassifikation: D4, D82, M4, M42, M48

Keywords: information economics, audit, auditor, credence good, efficiency.

Schlagwörter: Informationsökonomik, Jahresabschlussprüfung, Wirtschaftsprüfer, Vertrauensgut (Credence Good), Effizienzverluste.

I. Einleitung

Vertrauensgüter sind allgegenwärtig. Ob Zahnarzt, Autowerkstatt, Reparatur oder Taxifahrt in fremden Städten, in unserem Alltag haben wir sehr häufig mit Dienstleistungen zu tun, deren Qualität und Angemessenheit der Preisgestaltung wir als Normalverbraucher auch nach Erhalt und Bezahlung nicht einschätzen können. Obwohl in den Wirtschaftswissenschaften seit Jahrzehnten Informationsasymmetrien analysiert werden, wurden die speziellen Fragestellungen von Vertrauensgutmärkten erst relativ spät als Forschungsfeld erkannt und bearbeitet. Der Vertrauensgutcharakter vieler Dienstleistungen wird seither näher ana-

*Kontakt: Petra Ritzer-Angerer, ifo-Institut München, Poschingerstr.5, 81679 München, E-Mail: ritzer-angerer@ifo.de.

lysiert und es wird versucht, Folgerungen daraus abzuleiten. So war es auch bei der Jahresabschlussprüfung, die erstmals 2009 als Vertrauensgut betrachtet und diskutiert wurde.

Dieser Beitrag greift diese wissenschaftliche Diskussion um Jahresabschlussprüfung als Vertrauensgut auf. Nach einer kurzen theoretischen Einführung über Informationsökonomik und Wirtschaftsprüfungstheorie (Abschnitt II.), fasst Abschnitt III. die Kennzeichen von Vertrauensgütern im Allgemeinen und bezogen auf die Jahresabschlussprüfung im Speziellen zusammen und bejaht die Vertrauensguteigenschaften der Jahresabschlussprüfung im Ergebnis. Abschnitt IV. widmet sich entsprechend den wirtschaftspolitischen Folgerungen der Vertrauensgutforschung für die Wirtschaftsprüferhaftung. Im Ergebnis folgt ein Plädoyer für eine Verschärfung der Wirtschaftsprüferhaftung. Eine solche wurde von juristischer Seite schon häufig gefordert, hier liegt nun erstmals auch eine wirtschaftswissenschaftlich begründete Analyse und Bestätigung dieser Forderung vor. Abschnitt V. fasst die Ergebnisse kurz zusammen und gibt einen Ausblick.

II. Informationsökonomik und Wirtschaftsprüfungstheorie

Die Informationsökonomik beschäftigt sich mit den Kosten der Beschaffung und Bereitstellung von Informationen, deren Austausch zwischen den Akteuren und den Wirkungen unterschiedlicher Zugänge zu Informationen auf die Funktionsweise ökonomischer Systeme.¹ Dabei sind vollkommene Informationen keine Voraussetzung für effiziente Märkte – unter dynamischen Gesichtspunkten sind diese gar nicht erwünscht – und es gibt im marktwirtschaftlichen Selbststeuerungsmechanismus mannigfaltige Optionen Informationsmängeln entgegenzuwirken. Führen Informationsmängel allerdings zu Allokationsproblemen, weil Tauschvorgänge in Folge systematischer Entscheidungsfehler der Marktteilnehmer unterbleiben und deshalb gegebenenfalls sogar Marktversagen droht (Fritsch 2011, S. 254), kann dies wirtschaftspolitische Eingriffe rechtfertigen (Berg/Cassel/Hartwig 2007). Dieser Beitrag behandelt die Frage, ob solche gravierenden Informationsmängel in Bezug auf die Qualität von Jahresabschluss-

¹ Als Begründer der modernen Informationsökonomie gelten *George A. Akerlof*, *Michael Spence* und *Joseph E. Stiglitz*, die im Jahr 2001 dafür den Nobelpreis für Wirtschaftswissenschaften erhalten haben, vgl. *Stiglitz* (2002).

prüfungen bestehen, d. h. ein Vertrauensgut vorliegt und inwieweit wirtschaftspolitische Eingriffe auf dem Prüfungsmarkt diesem Problem entgegenwirken können.

Die Wirtschaftsprüfungstheorie basiert auf der Annahme einer asymmetrischen Informationsverteilung, welche als grundsätzliche Motivation von Wirtschaftsprüferleistungen dient (*Marten/Quick/Ruhnke* 2015, S. 48 f.) Die asymmetrische Informationsverteilung besteht zwischen informationsrechtlich privilegierter Unternehmensleitung und schlechter informierten Abschlussadressaten und begründet die fünf Funktionen der Jahresabschlussprüfung: Kontrollfunktion, Korrekturfunktion, Informationsfunktion, Beglaubigungsfunktion sowie Präventivfunktion (*Coenenberg/Haller/Schultze* 2018, S. 1003). Die Beglaubigungsfunktion wird dadurch wahrgenommen, dass der Abschlussprüfer einen Bestätigungsvermerk für den geprüften Abschluss erteilt. Der hier verwendete Terminus Beglaubigung, englisch *credence*, weist rein begrifflich bereits darauf hin, dass es sich bei der Jahresabschlussprüfung um ein Vertrauensgut (*credence good*) handeln könnte.

Die Vertrauensguttheorie ist ein Teilgebiet der Informationsökonomik und die Idee, sie auf die Jahresabschlussprüfung anzuwenden, stammt von *Monika Causholli* (2009). Dieser Beitrag überprüft Caushollis Überlegungen zur Rechtfertigung der Annahme, die Jahresabschlussprüfung sei ein Vertrauensgut und beleuchtet die Anwendung wissenschaftlicher Arbeiten zur Vertrauensguttheorie auf den Markt für Jahresabschlussprüfungen. Im Folgenden werden Vertrauensgüter zunächst definiert und geprüft, ob die Jahresabschlussprüfung die Kriterien dieser Definition erfüllt. Im Anschluss werden die Marktmechanismen und -bedingungen präsentiert, die als generelle Lösung der typischen Vertrauensgutproblematik gelten und diese Modellbetrachtungen auf den Markt für Jahresabschlussprüfungen übertragen.

III. Die Jahresabschlussprüfung als Vertrauensgut

Zunächst wird die Frage beantwortet, was Vertrauensgüter kennzeichnet. Anschließend wird geprüft, ob diese charakteristischen Merkmale bei der Jahresabschlussprüfung vorliegen.

1. Was kennzeichnet Vertrauensgüter?

Vertrauensgüter (credence goods) sind Waren bzw. Dienstleistungen, bei denen der Anbieter bzw. Experte mehr über die vom Kunden benötigte Qualität weiß als der nachfragende Kunde selbst (*Dulleck/Kerschbamer* 2006, S. 5). In der ursprünglichen Klassifikation von Gütern in normale Güter (ordinary goods), Suchgüter (search goods) und Erfahrungsgüter (experience goods) waren Vertrauensgüter (credence bzw. post-experience goods) noch nicht enthalten (*Nelson* 1970, S. 311 ff.) und wurden erst später ergänzt (*Darby/Karni* 1973, S. 67 ff.). Als typisches Lehrbuchbeispiel gelten Reparaturen, in der jüngeren Vergangenheit wurde im Zuge der Diskussion über eine neue Regulierung sowohl Taxifahrten² als auch Arzneimittel³ als Vertrauensgüter erörtert.

Der Informationsvorsprung des Anbieters gegenüber dem nachfragenden Kunden in Bezug auf die Qualität des angebotenen Gutes bzw. der Dienstleistung kann zu Negativauslese führen (adverse selection): Wenn der Kunde seine Zahlungsbereitschaft an seine durchschnittliche Qualitätserwartung koppelt und der Anteil minderwertiger Angebote ein kritisches Maß überschreitet, lohnt es sich für Anbieter qualitativ hochwertiger Waren bzw. Dienstleistungen nicht mehr, sie werden vom Markt gedrängt (bekannt als Market-for-lemons-Problem) (*Akerlof* 1970, S. 488). Dabei hängt das genaue Ausmaß der Informationsasymmetrie entscheidend von den Eigenschaften der angebotenen Artikel bzw. der angebotenen Dienstleistungen ab. So besteht zum Beispiel bei Gebrauchtwagen keine Informationsasymmetrie in Bezug auf Eigenschaften wie Marke und Farbe. Bei Gebrauchtwagen kann sich das Problem der Negativauslese aufgrund asymmetrischer Informationsverteilung so gravierend auswirken, dass es zum Zusammenbruch des Marktes führt. Asymmetrische Informationsverteilung ist dementsprechend kein binäres Merkmal, liegt also vor oder nicht, sondern entscheidend ist das konkrete Ausmaß des Informationsvorsprungs.

Bei Vertrauensgütern, die ja ausschließlich von Experten angeboten werden, liegt ein doppelter Informationsvorteil (Double-information-advantage (*Birchler/Büttler* 2007, S. 389)) vor, da sie nicht nur besser wissen, wie man ein Problem behandeln oder reparieren kann, sondern auch, welches Problem genau vorliegt. Im Vergleich zum einfachen Informationsvorsprung des Anbieters gegenüber dem Kunden entsteht damit ein Informationsvorsprung in Bezug auf Defektanalyse und in Bezug auf die beste Reparaturmaßnahme des speziellen Defekts. Sie

² Vgl. *Wambach* u. a. 2018, S. 114–128 (118). Regulierungsoptionen für Taximärkte wie Navigationsgeräte oder feste Fahrpreise für typischerweise von Fremden gebuchte Routen (wie z. B. Flughafen-Zentrum) wurden erst kürzlich diskutiert, vgl. *Kerschbamer/Sutter* 2017, S. 7.

³ Vgl. *Zimmerer* u. a. 2014, S. 6–37 (26 f.) mit weiteren Ausführungen.

verkaufen also Diagnose *und* zugehöriges Heilmittel. Durch diese Situation entstehen Anreize zur Unehrlichkeit (Birchler/Bütler 2007, S. 389). Werden Diagnose und Behandlung getrennt oder von zwei verschiedenen Experten durchgeführt, lassen sich dagegen stets First-Best-Lösungen erreichen (Bester/Dahm 2017; Bester/Ouyang 2018; Chen/Li/Zhang 2017).

a. Ausmaß des Informationsvorsprungs bestimmt strategisches Potential

Beobachtbare Eigenschaften der Ware bzw. Dienstleistung sind am geringsten von Informationsasymmetrien betroffen, da sie *vor* dem Kauf – also *ex ante* – in Erfahrung gebracht werden können (search attributes).⁴ Bei Erfahrungsattributen (experience attributes) ist dies schwieriger, da der Käufer *nur selbst nach dem Kauf*, also erst *bei Gebrauch*, eigene Erfahrungen mit der Ware oder der Dienstleistung machen kann.⁵ Die extremste Form der Informationsasymmetrie besteht bei Waren bzw. Dienstleistungen, deren Qualität selbst *nach dem Kauf* – d. h. *ex post* – nicht vom Käufer beurteilt werden kann (credence attributes) (Dulleck/Kerschbamer 2006, S. 5 ff.). Güter kombinieren in der Regel diese drei Kategorien von Eigenschaften mit unterschiedlicher Gewichtung, so dass jeweils unterschiedliche große Informationsvorsprünge und Güter mit unterschiedlich starken Vertrauensguteigenschaften resultieren (Gottschalk 2017 u. 2018, S. 5).

Die Zuordnung dieser Attribute zu Waren und Dienstleistungen dient der Einschätzung des Ausmaßes der Informationsasymmetrie zwischen Verkäufer und Käufer, welche wiederum das Potential des Verkäufers bestimmt, strategisch zu handeln und seinen Informationsvorsprung opportunistisch zu nutzen. Diese Gefahr ist bei beobachtbaren Eigenschaften (search attributes) gering, da der Käufer schon vor dem Kauf Möglichkeiten besitzt, Qualitätsprobleme in Verhandlung mit dem Verkäufer oder durch Rücktritt vom Kauf zu lösen. Bei Erfahrungseigenschaften (experience attributes) ist sie schon etwas höher, da erst nach dem Kauf das Problem erkannt wird, aber in zeitlicher Begrenzung zumindest noch Gewährleistungs- oder Garantieansprüche geltend gemacht werden können. Bei Vertrauenseigenschaften (credence attributes) ist die Gefahr strategischen Handelns zum eigenen Vorteil des Anbieters am ausgeprägtesten, da der Verkäufer

⁴ Die Unterscheidung von Eigenschaften in die drei Gruppen search, experience und credence geht auf Nelson 1970, S. 311 ff. und Darby/Karni 1973, S. 67 ff. zurück (siehe auch oben).

⁵ Dies stellt bei Geräten aller Art ein entscheidendes Problem dar. Die Hersteller kennen dieses Problem als Kaufhemmnis. Deshalb wird dem Kunden durch verschiedene Maßnahmen oft ermöglicht, Geräte erst einmal selbst auszuprobieren, zum Beispiel durch Testgeräte in Geschäften, Vorführungen der Geräte in der Werbung, etc.

daraufsetzen kann, dass der Käufer das Qualitätsproblem nie erkennen wird bzw. der Nachweis des Qualitätsproblems sehr hohe Kosten verursachen würde.

b. Strategisches Potential des Anbieters eines Vertrauensguts

Beim typischen, in der Literatur gern analysierten Beispiel für ein Vertrauensgut, der Autoreparatur, weiß der nachfragende Kunde zwar, dass bei seinem Fahrzeug ein Problem auftritt, da er aber nicht weiß, durch welchen Mangel das Problem verursacht wird, kann er nicht sagen, wie umfangreich die von ihm benötigte Reparatur tatsächlich ist. Das einzige, was er beobachten kann, ist, ob nach der Reparatur das Problem erneut auftritt.

Der Anbieter von Autoreparaturen hat grundsätzlich zwei Möglichkeiten, seinen Informationsvorsprung strategisch zu nutzen (*Dulleck/Kerschbamer* 2006, S. 7):

Falls der Kunde einen hochspeziellen, komplizierten und teuren Eingriff benötigt, er aber nur einen sehr einfachen, billigen erhält, spricht man von *Undertreatment*; umgekehrt kann bei einem einfach und schnell behebbaren Fehler sehr aufwändig vorgegangen werden, was als *Overtreatment* bezeichnet wird (*Dulleck/Kerschbamer* 2006, S. 5 ff.).⁶

Darüber hinaus gibt es bei einfachen und billigen Reparaturen stets die Möglichkeit, die simple Reparatur vorzunehmen und eine teure in Rechnung zu stellen, was in der Literatur *Overcharging* genannt wird. *Overtreatment* und *Overcharging* stellen klassische *Moral Hazard*-Probleme dar (*Birchler/Bütler* 2007, S. 371).

c. Empirische Belege

Empirische Studien enthalten klare und deutliche Hinweise für das strategische Verhalten von Anbietern von Vertrauensgütern (*Hilger* 2016). So hat eine Studie

⁶ Beide Strategien weichen vom Optimum, der maßgeschneiderten Behandlung des Problems ab. Welche der beiden Strategien attraktiv wird, hängt von der aktuellen Auftragslage und der vertraglichen Leistungsbeschreibung ab. Führt *Undertreatment* zur kurzfristigen Fehlerbehebung und ist eine solche vertraglich zugesichert, kann der Anbieter bei aktuell guter Auftragslage den Auftrag parken und auf später verschieben (wenn dann das unzureichend behandelte Problem erneut auftritt). Basiert die vertraglich zugesicherte Leistung darauf, dass die Entlohnung auf Stundenbasis erfolgt und ist zusätzlich die Auftragslage nicht gut, kann *Overtreatment* eine attraktive Option darstellen.

von 1993 (Wolinsky 1993, S. 380 ff.) gezeigt, dass mehr als die Hälfte der Autoreparaturen unnötig ist und eine von 1997, dass Ärzte selbst deutlich weniger chirurgische Eingriffe bekommen als Durchschnittspatienten, was auf Overtreatment hindeutet (Emons 1997, S. 107 ff.; vgl. auch Chiappori/Durand/Geoffard 1998, S. 499 ff.; Einav u. a. 2013, 178 ff.). Bei Autoreparaturen gibt es ebenfalls Studien mit sehr deutlichen Hinweisen auf Undertreatment (Schneider 2012, S. 406 ff.). Auch Anlage- und Rechtsberatung weist Eigenschaften von Vertrauensgütern auf (Bluethgen//Meyer/Hackethal 2008; Hadfield 2000, S. 953 ff.). Ebenso lässt sich strategisches Verhalten bei Taxifahrten belegen, welches im Auftrag von Fahrgästen, die mit dem Terrain unvertraut sind oder zumindest so wahrgenommen werden (wollen), nachweislich häufiger zu beobachten ist (Balafoutas/Kerschbamer/Sutter 2017, S. 1 ff.; Kerschbamer/Sutter 2017, S. 5 ff.).

Keinen Hinweis auf asymmetrische Information bzw. adverse Selektion finden allerdings Chiappori/Salanié (2000) auf dem Markt für KFZ-Versicherungen und Doran/Robertson/Henry (2005) auf dem Markt für verschreibungspflichtige Medikamente. Über die Rolle von Versicherungen oder Erstattungen an den Konsumenten wird diskutiert, es gibt Hinweise auf eine Verstärkung der Moral Hazard⁷-Problematik (Kerschbamer/Neururer/Sutter 2016), was als Second-degree-moral-hazard⁸ bezeichnet wird (Balafoutas/Kerschbamer/Sutter 2017; Kerschbamer/Sutter 2017, S. 16). Experimentell zeigt sich beispielhaft bei Computerreparaturen, dass Anbieter, die erfahren, dass die Reparaturkosten von einer Versicherung übernommen werden, im Durchschnitt bis zu 80 % höhere Beträge in Rechnung stellen (Balafoutas/Kerschbamer/Sutter 2017, S. 7457). Der Schaden für Versicherungen durch diesen Effekt ist erheblich. Besonders betroffen ist davon der Markt für Medikamente und medizinische Behandlungen (Lu 2014), zumal Krankenversicherungen international Standard sind. Für andere Vertrauensgüter wie zum Beispiel der Rechtsberatung existieren Versicherungsverträge wie Rechtsschutzpolicen. Der Vertrauensgutcharakter wird sogar bevorzugt als Argument für den Abschluss einer Versicherung verwendet.

⁷ Mit *Moral Hazard* wird der direkte Einfluss asymmetrischer Informationsverteilung auf das Marktgeschehen bzw. das Marktergebnis bezeichnet.

⁸ *Second-degree-moral-hazard-Effekte* wurden insbesondere auf regulierten Märkten untersucht. Der Begriff enthält der direkten *Moral Hazard* (s. o.) nachgeschaltete Effekte wie die Tendenz der Anbieter eines Marktes auf durch die Nachfrager antizipierte Moral Hazard-Probleme zu reagieren.

2. Erfüllt die Jahresabschlussprüfung die Kennzeichen eines Vertrauensguts?

a. Informationsvorsprungs des Wirtschaftsprüfers

In der wissenschaftlichen Literatur, die sich mit der Frage befasst, ob die Abschlussprüfung Eigenschaften eines Vertrauensguts aufweist, werden insbesondere zwei Aspekte hervorgehoben (*Causholli/Knechel* 2012, S. 634 ff.):

Das Ergebnis einer Jahresabschlussprüfung ist *nicht beobachtbar* in dem Sinn,

- dass das Restrisiko eines unentdeckten Fehlers grundsätzlich nicht ausgeschlossen werden kann und
- dass das im Einzelfall tatsächlich erreichte Sicherheitsniveau niemals bekannt wird (*Causholli/Knechel* 2012, S. 634 f.; *Böcking/Gros* 2015, S. 307).

Aufgrund der speziellen und unsicheren Natur von Prüfungsprozessen kann allein der durchführende Prüfer entscheiden, wieviel Anstrengung aufgewendet und belegt werden muss, um eine Gewährleistung der Prüfungsstandards zu erreichen. Als Hinweis für seine Einschätzung des Arbeitsaufwands kann das für die Prüfung veranschlagte Honorar herangezogen werden. Die Beurteilung des Prüfungsumfangs hat ex ante zu erfolgen, ist jedoch selbst ex post schwierig. Das zu prüfende Unternehmen kann zu keinem Zeitpunkt sicher sein, das ursprünglich in Aussicht gestellte Sicherheitsniveau auch zu bekommen (*Causholli/Knechel* 2012, S. 634 f.; *O’Keefe/Simunic/Stein* 1994, S. 241 ff.; *Knechel/Rouse/Schelleman* 2009, S. 1607 ff.).

Der Informationsvorteil des Abschlussprüfers erstreckt sich auf Planung und Durchführung der Prüfung (*Causholli/Knechel* 2012, S. 635). Abhängig von der Wahrnehmung des zu prüfenden Unternehmens durch den Prüfer, wirkt sich der Informationsvorteil auf geplante strategische Entscheidungen im Prüfungsprozess aus. Das Management hat zwar ex ante einen Informationsvorsprung in Bezug auf das Unternehmen und seine Abläufe, der Prüfer hat seinerseits einen Informationsvorsprung bezüglich der geplanten Prüfungsdurchführung. Dieser Informationsvorteil bezüglich der Prüfungsdurchführung wird im Verlauf der Prüfung immer größer, da erste Prüfungsergebnisse vorliegen, die wiederum Entscheidungen nach sich ziehen, deren Ergebnisse später wiederum weiteren Prüfungsentscheidungen zu Grunde liegen. Der Prüfer durchläuft einen Lernprozess, der seinen Informationsnachteil in Bezug auf die Unternehmensabläufe reduziert, gleichzeitig wächst sein Informationsvorsprung in Bezug auf die Details der Prüfungsdurchführung. Nach Abschluss der Prüfung und Anfertigung des Prüfungsberichts fügen sich die Erfahrungen des Prüfers in den einzelnen Prü-

fungsschritten zu einem Bild des Unternehmens zusammen. Die bei der Planung getroffenen Annahmen, auch die, welche der angebotenen Honorarfestsetzung zugrunde lagen, können korrigiert werden.

b. Übertragung auf Kunde-Experte-Beziehung

Es existiert also ein Informationsvorteil des Abschlussprüfers, der sich auf Planung und Durchführung der Prüfung erstreckt. Anders als bei anderen Vertrauensgütern, bei denen der Kunde typischerweise selbst überhaupt nicht weiß bzw. wissen kann, welche z. B. Reparatur er benötigt, kann bei der Abschlussprüfung davon ausgegangen werden, dass das Management des zu prüfenden Unternehmens selbst einschätzen kann, ob es ein hohes oder niedriges Risiko aufweist (*Causholli/Knechel* 2012, S. 640, 645; *Knechel/Willekens* 2006, S. 1344 ff.). Weder das Management, das als Ansprechpartner für die Jahresabschlussprüfer während der Prüfungsdurchführung dient, noch der Aufsichtsrat stellen jedoch den Konsumenten im Vertrauensgütermodell dar, einen möglichen Schaden durch die genannten Strategien erleiden die Eigentümer.

Im Vergleich zu einem typischen Vertrauensgutanbieter ist die Situation bei der Jahresabschlussprüfung komplizierter, weil der Käufer der Abschlussprüfung und der letztendliche Konsument im Modell nicht ohne weiteres identisch sind.⁹ Der Wirtschaftsprüfer wird zwar von den Gesellschaftern bestellt (§ 318 HGB), der Aufsichtsrat schlägt jedoch einen Wirtschaftsprüfer vor und dieser Vorschlag wird in der Haupt- bzw. Gesellschafterversammlung in der Regel auch angenommen.¹⁰ Faktisch gibt damit in der Regel der Aufsichtsrat die Jahresabschlussprüfung in Auftrag, die Gesellschaftermehrheit verfügt jedoch über einen Zustimmungsvorbehalt. Bei der GmbH und haftungsbeschränkten Personenhandelsgesellschaften (§ 264a HGB) kann das Recht den Wirtschaftsprüfer zu wählen, entsprechend einer abweichenden Regelung im Gesellschaftsvertrag ganz auf einen Aufsichtsrat, einen Beirat oder einen Gesellschafterausschuss übertragen werden.¹¹ Unter bestimmten Umständen kann gem. § 318 Abs. 3 u. 4 HGB der Abschlussprüfer

⁹ Das Management fungiere als Intermediär, vgl. *Causholli/Knechel* 2012, S. 642f.

¹⁰ Bei Aktiengesellschaften bestimmt § 124 Abs. 3 S. 1 AktG: „Zu jedem Gegenstand der Tagesordnung, über den die Hauptversammlung beschließen soll, haben der Vorstand und der Aufsichtsrat, zur Wahl von Aufsichtsratsmitgliedern und Prüfern nur der Aufsichtsrat, in der Bekanntmachung Vorschläge zur Beschlussfassung zu machen.“

¹¹ Bei Gesellschaften mit beschränkter Haftung und bei offenen Handelsgesellschaften sowie Kommanditgesellschaften (i. S. v § 264a Abs. 1 HGB) kann der Gesellschaftsvertrag etwas anderes bestimmen (§ 318 Abs. 1 S. 2 HGB).

durch ein Gericht bestellt werden, wenn dies vom gesetzlichen Vertreter, vom Aufsichtsrat oder von den Gesellschaftern beantragt worden ist (Coenenberg/Haller/Schultze 2018, S. 999).

Überträgt man das Kunde-Experte-Vertrauensgutmodell auf die Jahresabschlussprüfung, stellen die Anteilseigner des geprüften Unternehmens in ihrer Gesamtheit den Kunden im Modell dar, denn es ist ihr Unternehmen, das geprüft wird. Der Wirtschaftsprüfer als Experte im Modell hat in der Regel keinerlei Kontakt zu den Eigentümern bzw. Kunden, sondern hat bei Durchführung der Prüfung nur mit dem Management zu tun, das zu den Anteilseignern als Kunden in einer Prinzipal-Agent-Beziehung steht. Lediglich auf der Hauptversammlung haben die Anteilseigner überhaupt die Möglichkeit Fragen zu stellen. Dieser mangelnde Kontakt zu den Eigentümern, dafür intensive Kontakt zu den Managern des Unternehmens birgt die Gefahr der *vorsätzlich* falschen Prüfungsdurchführung, wobei einige Stimmen in der Literatur sogar davon ausgehen, dass Bilanzfälschungen nicht selten durch Prüfer geduldet bzw. in Unterstützung des Managements mit diesem gemeinsam betrieben werden (Müller 2004, S. 217; Bauer 2017, S. 76).

Damit wird die Kernfunktion der Jahresabschlussprüfung, die Überwachung der Arbeit des Managements, unterlaufen. Eigentumsschäden können den Anteilseignern vor allem durch wertvernichtende Investitions- und Finanzierungsentscheidungen der Manager (Nichtrealisierung vorteilhafter und Realisierung suboptimaler Investitionsprojekte) (Drukarczyk 1993, S. 305 ff.) entstehen. Die Jahresabschlussprüfung stellt schon konzeptionell keine Prüfung der Geschäftsführung in Bezug auf Werteffekte getroffener Investitions- und Finanzierungsentscheidungen dar.¹²

In Bezug auf die Übertragung der Kunde-Experte-Beziehung des Vertrauensgutmodells auf die Jahresabschlussprüfung bedeutet das, dass es sich hier um eine *verschärfte* Vertrauensgutproblematik handelt. Der Abschlussprüfer hat nicht nur die für Vertrauensgüter typischen Möglichkeiten, seinen Informationsvorsprung opportunistisch auszunutzen, er arbeitet bei der Erbringung seiner Dienstleistung eng mit den Managern zusammen, die zu den Eigentümern (im Vertrauensgutmodell Kunden) ebenfalls in einer Prinzipal-Agent-Beziehung stehen. Aufgrund ihrer Prinzipal-Agenten-Beziehung zu den Eigentümern haben Manager Anreize, ihre Erfolge herauszustellen und ihre Misserfolge zu verbergen. Dabei ist die Wahrscheinlichkeit, dass Abschlussprüfer das in sie von

¹² Das OLG Düsseldorf hat dies so formuliert, dass die Abschlussprüfung „keine Kreditwürdigkeits-, Organisations-, Steuer-, Rentabilitätsprüfung oder Prüfung der Geschäftsführung auf kaufmännische Vernünftigkeit“ ist, Urteil vom 19.11.1998, 8 U 59/98, Rn. 98, juris.

der Öffentlichkeit gesetzte Vertrauen verletzen deutlich höher als die, dass sie ihre Beziehung zum Auftraggeber gefährden (Neu 1991). In der Konstellation der Kooperation zweier Agenten desselben Prinzipals entsteht aus der Zusammenarbeit schnell ein Zusammenhalt¹³, zumal beide Agenten keinen Kontakt zu ihrem Prinzipal haben.

In der Vertrauensgutliteratur wurden die Kontrollmöglichkeiten des Kunden stets besonders hervorgehoben, obwohl sie gering sind: Bei der Autoreparatur sollte der Kunde gezielte Fragen stellen und sich das defekte und ausgetauschte Ersatzteil immer aushändigen lassen. Natürlich kann der Kunde nicht sichergehen, dass er ein Teil erhält, das in seinem Auto tatsächlich verbaut war, aber er macht es damit dem Anbieter zumindest schwerer, ihn zu übervorteilen. Diese Kontaktmöglichkeiten zum Aufbau von Rechtfertigungsdruck sind bei der Jahresabschlussprüfung weitgehend unterbunden, was ebenfalls die Vertrauensgutproblematik verstärkt.

c. Strategisches Potential bei Jahresabschlussprüfungen

Die Frage, ob der Jahresabschlussprüfer über ein vergleichbares strategisches Potential verfügt wie der Anbieter eines typischen Vertrauensguts, ist damit schon zum Teil vorweggenommen. Analysiert man die beiden Möglichkeiten, den Informationsvorsprung auszunutzen, getrennt, gilt folgendes: Grundsätzlich besteht die Möglichkeit von Overtreatment, indem der Wirtschaftsprüfer zusätzliche, zur Erfüllung der Standards gar nicht erforderliche Prüfungen vornimmt. Allerdings ist dieses strategische Potential durch das ursprüngliche Angebot begrenzt. Den Rahmen des Angebots übersteigende Prüfungsleistungen müssten mit dem Auftraggeber besprochen und gerechtfertigt werden, was ein klares Hemmnis darstellt.¹⁴ Der Möglichkeit des Undertreatment, das heißt, Prüfungsleistungen zu unterlassen, die die Einhaltung der Standards erfordern, stellen Verletzungen der Berufspflichten (§ 323 HGB, § 43 WPO) insbesondere der Gewissenhaftigkeit dar, sodass diese vom Jahresabschlussprüfer auf jedem Fall durchzuführen sind, selbst wenn sie den Rahmen des ursprünglichen Angebots

¹³ Zum Problem der „persönlichen Vertrautheit“ und deren Folgen für die Objektivität des Prüfers, vgl. Bauer 2017, S. 86 f. mit weiteren Nachweisen.

¹⁴ Gemäß Berufssatzung ist eine angemessene Vergütung erforderlich (§ 43 Abs. 1 S. 2 Berufssatzung): „Ein Pauschalhonorar darf für einen Prüfungs- oder Gutachtenauftrag grundsätzlich nur vereinbart werden, wenn festgelegt wird, dass bei Eintritt nicht vorhersehbarer Umstände im Bereich des Auftraggebers, die zu einer erheblichen Erhöhung des Aufwandes führen, das Honorar entsprechend zu erhöhen ist.“ (§ 43 Abs. 2 Berufssatzung).

sprengen. Auch *Causholli/Knechel* (2012) sehen daher eher die Gefahr von Overtreatment,¹⁵ nicht zuletzt weil sie dem Gewinnstreben der Wirtschaftsprüfungsgesellschaft entspricht (*Causholli* 2009; *Causholli/Knechel* 2009), in der Praxis könne jedoch auch Undertreatment nicht völlig ausgeschlossen werden, insbesondere wenn Zeit- und Budgetdruck im Einzelfall besonders groß sind (*Causholli/Knechel* 2012, S. 641).

Overcharging, die zweite strategische Möglichkeit, ist bei Jahresabschlussprüfungen begrenzt möglich, da der Prüfungsbericht (§ 321 HGB) einen eindeutigen Zusammenhang zwischen Prüfungsleistung und dafür in Anspruch genommener Zeit herstellt. Zusätzlich kann ein sorgfältiger Vergleich mit dem Ursprungsangebot vorgenommen werden. Der Auftraggeber kann in jedem Fall eine Plausibilitätskontrolle durchführen, die umso detaillierter möglich ist, je mehr Informationen ihm über den Prüfungsablauf zur Verfügung stehen.¹⁶

3. Die Jahresabschlussprüfung erfüllt Kriterien eines Vertrauensguts

Die Jahresabschlussprüfung erfüllt nicht nur Vertrauensguteigenschaften, es liegt eine verschärfte Vertrauensgutproblematik vor. Dies ist bedingt durch zweierlei: Die Jahresabschlussprüfung hat konzeptionell gar nicht die Funktion, wertbeeinflussende Investitions- und Finanzierungsentscheidungen des Managements zu überprüfen, entscheidend sind der handelsrechtlich genau definierte Prüfungsumfang (§ 317 HGB) und die vom Institut der Wirtschaftsprüfer (IDW) veröffentlichten Prüfungsstandards, die als Grundsätze ordnungsgemäßer Abschlussprüfung (GoA) gesetzliche Regelungen ergänzen. Zweitens arbeitet der Abschlussprüfer als Vertrauensgutanbieter mit dem Management zusammen, das selbst – wie der Abschlussprüfer – als Agent einer Prinzipal-Agent-Beziehung zu den Eigentümern tätig ist.

Dieser Situation wird haftungsrechtlich durchaus bereits Rechnung getragen: Zunächst ist es Aufgabe und Verantwortung der zuständigen Organe der bilanzierenden Gesellschaft, dafür zu sorgen, dass die Rechnungslegung den gesetzlichen

¹⁵ Vgl. *Causholli/Knechel* 2012, S. 641 mit Hinweisen, dass es bereits Beschwerden über Overtreatment gab, wobei Overtreatment mit den hohen Anforderungen des PCAOB an Prüfungen von Seiten der Wirtschaftsprüfer gerechtfertigt wurde. Allerdings bekräftigen dieselben Autoren auch die Gefahr von Under-Auditing, vgl. *Causholli/Knechel/Lin/Sappington* 2013, S. 573–605.

¹⁶ *Causholli/Knechel* (2012) weisen darauf hin, dass Overcharging bei Jahresabschlussprüfungen simpel das Abrechnen nicht gearbeiteter Stunden bedeutet bzw. – etwas subtiler – Stundensätze von höher qualifiziertem Personal als dem tatsächlich eingesetzten anzusetzen, S. 642.

Vorgaben entspricht, was haftungsbewehrt ist (§ 93 Abs. 2 AktG für Aktiengesellschaften). Nach der Prüfung durch den Jahresabschlussprüfer (§§ 316 ff. AktG) ist dann noch einmal der Aufsichtsrat mit der Prüfung der Abschlüsse beauftragt (§ 171 AktG), diese Prüfpflicht ist haftungsbewehrt (§ 116 AktG). In Deutschland ist ebenfalls das Konstrukt des Mitverschuldenseinwands gesetzlich verankert (*Doralt* 2015, S. 269 ff.), das grundsätzlich bei jedem Haftungsfall zu prüfen ist. Trotz dieser Vielzahl von Akteuren, die für die Erstellung und Prüfung von Abschlüssen zuständig sind, geraten insbesondere die Abschlussprüfer in der öffentlich wahrgenommenen Diskussion in Kritik (*CICA* 1988, S. 15¹⁷), was als ungerechtfertigt erstaunt (*Hennrichs* 2015, S. 249) und in der wissenschaftlichen Auseinandersetzung mit der Wirtschaftsprüferhaftung immanent mitschwingt und eine sachliche Auseinandersetzung erschwert.

Bei der Diskussion der Regulierung der Wirtschaftsprüferhaftung (§ 323 HGB) ist stets dieser haftungsrechtliche Kontext zu beachten. Die aktuellen europäischen Regime der Wirtschaftsprüferhaftung sind Folge der EU-Empfehlung zur Beschränkung der zivilrechtlichen Haftung von Abschlussprüfern und Prüfungsgesellschaften aus dem Jahr 2008 (Amtsblatt der Europäischen Union vom 21.6.2008, L 162/39 u. 40).

Prüfungsqualität kann darüber definiert werden, wie gut das angemessene Ausmaß an zu erbringenden Prüfungsleistungen dem Ausmaß entspricht, das der Prüfer vorgibt und einhält, wobei jede Abweichung vom Idealbild der angemessenen Prüfungsleistung einer Qualitätsreduktion entspricht (*Causholli/Knechel* 2012, S. 642). Die Jahresabschlussprüfung wird zur Gruppe der typischen Vertrauensgüter gezählt, da die Anbieter Expert providers darstellen (*Gottschalk* 2017 und 2018, S. 4 und 6), d. h. sie zählt zu den Gütern mit relativ starken Vertrauensguteigenschaften.

Zusammenfassend lässt sich damit sagen, dass es sich bei der Jahresabschlussprüfung um ein Vertrauensgut handelt, wobei die Konzeption der Jahresabschlussprüfung und das bestehende Prinzipal-Agent-Geflecht, in das der Abschlussprüfer bei Prüfungsdurchführung eintritt, die Vertrauensgutproblematik verschärfen. Grundsätzlich sind alle drei typischen Strategien von Vertrauensgutanbietern in entsprechenden Prüfungssituationen denkbar. Die wenigen Kontrollmöglichkeiten des Kunden im Vertrauensgutmodell werden bei der Abschlussprüfung weitgehend unterbunden. Das hier erzielte Ergebnis, dass es sich bei der Jahresabschlussprüfung um ein Vertrauensgut handelt, wird von

¹⁷ Dort heißt es „It almost seems as though the public equates a business failure with an audit failure“, *CICA*, Report of the commission to study the public's expectations of audits, 1988, S. 15, Tz. 2.18.

zahlreichen Stimmen in der Literatur geteilt (*Baetge/Matena*, S. 185; *Ballwieser* 2008, S. 4 f.; *Bauer* 2017, 52; *Eisenhardt* 2011, S. 119; *Fleischer* 2002, S. F 23; *Herkendell* 2007, S. 98; *Kremer* 2007, S. 135; *Sattler* 2011, S. 68; *Wiemann* 2011, S. 42; *Wild/Scheithauer* 2012, S. 188).

IV. Wirtschaftspolitische Folgerungen aus der Vertrauensgutforschung für die Wirtschaftsprüferhaftung

Bei Vertrauensgütern sind die vielfältigen bekannten Methoden der Agency-Theorie zur Abmilderung von Problemen durch asymmetrische Informationsverteilung nicht mehr unbedingt wirksam. So können sich hohe gesetzliche und ethische Standards oder Reputation als glaubwürdiges Signal für qualitätvolle Arbeit, etc. als unwirksam erweisen (*Bonroy/Constantatos* 2008, S. 239). Dieser Beitrag konzentriert sich – basierend auf der Annahme, die Jahresabschlussprüfung erfülle Vertrauensguteigenschaften (s. o.) – auf die Marktmechanismen, die für Situationen mit extremer Informationsasymmetrie wie sie für Vertrauensgüter kennzeichnend ist, als wirksam gelten.

1. Marktmechanismen und Effizienzverluste auf Märkten für Vertrauensgüter

a. Grundmodell von *Dulleck/Kerschbamer* (2006)

Welche der identifizierten Anbieterstrategien auf den Märkten für Vertrauensgüter jeweils Effizienzverluste verursacht, wird stark von den jeweiligen Konditionen beeinflusst, die den rechtlichen Rahmen für die jeweiligen Markttransaktionen bilden. Die Abrechnung überhöhter Forderungen (Overcharging) zum Beispiel ist nur möglich, wenn der Kunde nicht überprüfen kann, welchen Service er tatsächlich erhalten hat.

Dulleck/Kerschbamer (2006) zeigen im Rahmen eines einfachen, vereinheitlichenden Modells, dass Märkte und ihre Institutionen unerwünschtes Anbieterverhalten kostenlos unterbinden, wenn drei kritische Annahmen über Marktmechanismen gelten und geben damit Aufschluss über die ideale Struktur von Märkten für Vertrauensgüter:

- **Homogenität** der nachfragenden Kunden ist gegeben;

- Verbundeffekte zwischen Diagnose und Eingriff (**Commitment**): Sobald der anbietende Experte seine Diagnose erstellt hat, sind Kunde und Experte auf vorgegebene Art und Weise verpflichtet, durch einen Eingriff fortzufahren¹⁸;
- Nachprüfbarkeit der Art des Eingriffs (**Verifiability**) und/oder eine Haftungsregelung (**Liability**) existiert, die den Kunden davor schützt, eine unsachgemäße und überteuerte Behandlung bezahlen zu müssen (nur eine der beiden Annahmen muss erfüllt sein) (Dulleck/Kerschbamer 2006, S. 8 ff.).

Im Modell wird Folgendes angenommen: Es gibt nur zwei Behandlungen, eine billige und eine teure. Der Kunde weiß, dass er ein Problem hat, weiß aber nicht, welche der beiden Behandlungen er braucht. Für alle Kunden gilt, dass mit Wahrscheinlichkeit h das große Problem und mit Wahrscheinlichkeit $(1-h)$ das kleine Problem vorliegt, so dass jeder Kunde gleich bewertet wird (Homogenität). Der Experte kann diagnostizieren, wie ernst das Problem ist und die jeweils geeignete Behandlung vornehmen und abrechnen oder er kann die Informationsasymmetrie opportunistisch ausnutzen und den Kunden übervorteilen. Sobald ein Behandlungsvorschlag des Experten vorliegt, ist der Kunde verpflichtet, die Behandlung bei ihm durchführen zu lassen (Commitment).¹⁹ Liegt das große Problem vor, kann der Experte nicht die billige Behandlung anbieten (Liability) und er kann auch nicht die teure abrechnen, wenn er die billige durchgeführt hat (Verifiability).

In drei von vier möglichen Fällen ist der Kunde zufrieden, da er in zwei Fällen ja den angemessenen Service erhält und im dritten Fall eine Behandlung erhält, die mindestens so gut ist wie die, die er tatsächlich benötigt (kleines Problem, teure Behandlung). Dies nimmt der Kunde annahmegemäß nicht wahr, da er ja nicht wissen kann, dass er nur das kleine Problem hatte und ein billigerer Eingriff das Problem ebenso hätte lösen können. So ist der Kunde im Ergebnis in drei der untersuchten vier Fälle zufrieden. Einzig im vierten Fall (großes Problem, billige Behandlung) wird er ex post feststellen, dass er eine andere Behandlung

¹⁸ Wenn es kostenlos möglich ist, eine Zweit- oder Drittmeinung einzuholen, kann der Anbieter nur durch öffentliche Preise Kunden anziehen. Bietet er den identischen Preis für die teure und die billige Behandlung an, würden stets alle Kunden die angemessene Behandlung erhalten. Gilt zusätzlich, dass es für den Kunden überprüfbar ist, welche Behandlung vorgenommen wurde (Verifiability), wird auch stets korrekt abgerechnet. Ironischerweise wäre das Einholen einer zweiten Meinung dann unnötig (Birchler/Bütler 2007, S. 392).

¹⁹ Dies ist eine Prämisse, die von den Autoren als durch den Vertrauensgutcharakter vorgegeben angenommen wird. Durch eine Trennung von Diagnose und Behandlung lassen sich First-Best-Lösungen erreichen (siehe 2.1). Die Commitment-Annahme ist deshalb also Teil des zu lösenden Problems, nicht Teil der Lösung.

gebraucht hätte als die, die er bekommen hat. Meistens weil das Problem erneut auftritt, so dass er erfährt, dass die erhaltene Behandlung falsch oder nicht ausreichend gewesen ist.

Dulleck/Kerschbamer (2006) analysieren und erörtern die auftretenden Effizienzverluste, wenn mindestens eine der Bedingungen nicht erfüllt ist. Dabei unterscheiden sie sieben Fälle mit unterschiedlichen Gleichgewichtseigenschaften, lediglich wenn sowohl Nachprüfbarkeit der Art des Eingriffs (Verifiability) als auch eine Haftungsregelung (Liability) fehlt (Fall 7), besteht die Gefahr, dass der Markt zusammenbricht. Daher muss im Modell nur eine dieser Bedingungen erfüllt sein (s. o.).²⁰

b. Unterschiedliche Auswirkungen einzelner Bedingungen

Ob diese beiden Bedingungen, von denen zur Erreichung von Markteffizienz optional nur eine erfüllt sein muss, gleich gut wirken, untersuchen *Dulleck/Kerschbamer/Sutter* (2011) in einer Fortführung ihrer Arbeit. Dazu erweitern sie ihr ursprüngliches Modell: Sie gehen dabei davon aus, dass die beiden Bedingungen Liability und Verifiability institutionell gewährleistet werden können und nehmen zusätzlich zwei neue Merkmale auf, nämlich Reputationsbildung (Reputation), d. h. der Kunde kann seinen Anbieter identifizieren (kein anonymer Markt) und Anbieterwettbewerb (Competition), d. h. es gibt die Option des Kunden zwischen verschiedenen Anbietern wählen zu können.

Anbieterhaftung (Liability) ist definiert wie im Grundmodell und schützt den Kunden wirksam vor Undertreatment. Auch die Nachprüfbarkeit der Art des Eingriffs (Verifiability) ist definiert wie im Grundmodell und schützt stets vor Overcharging. Da die beiden Bedingungen Liability und Verifiability jeweils erfüllt und nicht erfüllt sein können, ergeben sich durch Permutation erneut vier verschiedene Situationen, die unterschiedliche institutionell bedingte Handlungsoptionen für die Anbieter beinhalten. Sind die zwei Bedingungen nicht erfüllt, ist der Anbieter in der Wahl seiner Handlung in Bezug auf Eingriff und Abrechnungsbetrag völlig frei (hier besteht die Gefahr eines Marktzusammenbruchs, siehe oben), sind beide Bedingungen dagegen erfüllt, muss er nicht nur den richtigen Eingriff vornehmen, sondern dafür auch den angemessenen Preis abrechnen. Ist nur Liability gegeben, muss er den richtigen Eingriff vornehmen, kann dafür aber zu viel abrechnen (d. h. Overcharging). Ist dagegen ausschließlich Verifiability

²⁰ Kritisch zur Wirksamkeit dieser Bedingungen *Li/Ouyang* (2016): Effizienz werde selbst unter Annahme von Liability und Verifiability nicht erreicht.

erfüllt, kann er frei den Eingriff wählen, muss dann jedoch den zugehörigen Preis abrechnen. Nur wenn beide Bedingungen erfüllt sind, stehen keine Optionen mehr zur Verfügung, die eine Schädigung des Kunden ermöglichen, sind dagegen beide Bedingungen nicht erfüllt, droht Marktversagen. Die Aussage des Grundmodells, dass nur eine der beiden Bedingungen erfüllt sein muss, damit Markteffizienz erreicht wird, wird daher hier von den Autoren näher spezifiziert bzw. korrigiert.²¹

In einem zweiten Schritt gehen *Dulleck/Kerschbamer/Sutter* (2011) mit den beiden neuen Merkmalen Reputationsbildung und Anbieterwettbewerb genauso vor und bilden durch Permutation wieder vier Fälle: Besteht kein Wettbewerb und keine Möglichkeit frühere Anbieter zu identifizieren, bilden sich Kunde-Anbieter-Paare rein zufällig und ein früherer Anbieter kann nicht von einer guten Reputation profitieren bzw. für eine schlechte Erfahrung durch Auftragsentzug sanktioniert werden (Erfahrungen anderer Kunden mit diesem Anbieter werden annahmegemäß nicht bekannt). Sind beide Bedingungen dagegen erfüllt, können Kunden ihren Anbieter identifizieren und dementsprechend gezielt zwischen den Anbietern wählen. Ist nur Reputationsbildung möglich, kann der Kunde seinen Anbieter zwar identifizieren, jedoch nicht zwischen verschiedenen Anbietern wählen, er bleibt also an diesen einen gebunden. Im umgekehrten Fall kann der Kunde entsprechend wählen, aber seinen Anbieter nicht identifizieren.

Kombiniert man die insgesamt auf diese Weise gebildeten Fälle – jeden der ersten vier Unterfällen mit jeden der zweiten vier Unterfälle – ergeben sich insgesamt 16 Fälle. Für die 16 Einzelsituationen stellen *Dulleck/Kerschbamer/Sutter* (2011) Vorhersagen über das Verhalten von Anbietern und Kunden auf, und überprüfen diese experimentell unter Einsatz von 936 Studenten und 16 Runden mit dem Ziel, Aussagen über die Effekte der vier Merkmale zu treffen.

Entgegen ihrer Vorhersagen bot bei fehlender Haftungsregel rund ein Viertel der Anbieter (27 %) trotzdem über alle Runden hinweg stets den passenden Eingriff an, waren also in allen Situationen ehrlich. Dies kann als verhaltensbedingte Besonderheit interpretiert werden und zeigt die Bedeutung sozialer Werte für das Verhalten auf. Je mehr Anbieter sich ohne wirtschaftspolitische Regulierung ehrlich verhalten, umso weniger ist eine entsprechende Regulierung erforderlich. Überraschenderweise zeigt das Merkmal Verifiability, also die Nachprüfbarkeit

²¹ Für den hier strittigen Fall, dass Verifiability erfüllt ist und Liability nicht, wurde im Grundmodell angenommen, dass die Anbieter es dann vorziehen, einen Einheitspreis für beide Eingriffe („equal mark-up prices“) zu bilden und anzubieten (*Dulleck/Kerschbamer* 2006, S. 15). Dies führt aufgrund der Commitment-Bedingung dazu, dass stets der richtige Eingriff vorgenommen wird und aufgrund der Verifiability-Bedingung auch abgerechnet wird (*Birchler/Bütler* 2007, S. 392).

des Eingriffs durch den Kunden, kaum Einfluss. Das erstaunt, weil Nachprüfbarkeit der Art des Eingriffs die asymmetrische Informationsverteilung klar reduziert. Allerdings kann bedingt durch den doppelten Informationsvorteil des Anbieters nur Overcharging, jedoch nicht Over- oder Undertreatment unterbunden werden. Anbieterhaftung (Liability) zeigt dagegen sowohl für das Stattfinden von Markttransaktionen als auch für die realisierte Markteffizienz großen Einfluss. Das lässt sich dadurch erklären, dass es für Kunden bei Anbieterhaftung attraktiv wird mit Anbietern ins Geschäft zu kommen, da Undertreatment ausgeschlossen werden kann (Dulleck/Kerschbamer/Sutter 2011, S. 546, 548). Diesem Vorteil des Kunden kommt große Bedeutung zu, da sich das Fehlen einer solchen Schutzvorrichtung sehr kostenintensiv auf die Kunden auswirkt, wenn der Eingriff gegebenenfalls mehrmals beauftragt und bezahlt werden muss. Dazu kommt, dass ausschließlich Undertreatment für den Kunden als Ärgernis wahrgenommen wird, weil sich nur in diesem Fall herausstellt, dass das Problem nicht hinreichend behandelt wurde, wohingegen bei Overtreatment der Kunde niemals erfährt, dass der von ihm bezahlte Preis zu hoch war (s. o.). Zudem hat sich Overtreatment im Experiment als nicht besonders relevant und Overcharging als harmlos (nur geringe Preisunterschiede) herausgestellt, was die relative Bedeutung von Undertreatment und dem Schutz davor zusätzlich verstärkt (Dulleck/Kerschbamer/Sutter 2011, S. 548 f.).

Das Merkmal Wettbewerb (Competition) hat sich ebenfalls als sehr einflussreich gezeigt, zum einen, um das Preisniveau zu reduzieren und zum zweiten, um das Handelsvolumen des Marktes anzukurbeln. Das Effizienzniveau konnte nur dann durch Wettbewerb gesteigert werden, wenn zusätzlich Überprüfbarkeit des Eingriffs (Verifiability) erfüllt war (Dulleck/Kerschbamer/Sutter 2011, S. 549.).

Das Merkmal Reputation war nur in einem Fall einflussreich in dem Sinne, dass der Handel angekurbelt wurde und zwar in dem Fall, dass sowohl Haftungsregeln als auch Wettbewerb, also die beiden einflussreichsten Merkmale fehlten (Dulleck/Kerschbamer/Sutter 2011, S. 549.). Haftungsregeln sind daher der dominante Marktmechanismus (Dulleck/Kerschbamer/Sutter 2011, S. 545).

c. Wissenschaftliche Diskussion

Zusammenfassen lässt sich dies wie folgt: Während das theoretische Grundmodell von Dulleck/Kerschbamer (2006) eindeutige theoretische Bedingungen formuliert, die Markteffizienz zur Folge haben, zeigen ihre späteren experimentellen Modellerweiterungen (2011), dass die Einführung von Nachprüfbarkeit des Eingriffs (Verifiability) kaum eine Effizienzverbesserung bringt und dass es vor allem auf Haftungsregeln ankommt. Wambach u. a. (2018) folgern daraus zum einen,

dass hier „verhaltensökonomische Besonderheiten“ vorliegen, die zu diesem „Widerspruch zwischen den Ergebnissen von theoretischen und experimentellen Studien“ führen und zum anderen, dass damit die „Anwendbarkeit der Ergebnisse der Vertrauensgüterforschung für Politikempfehlungen ungeeignet werden“ (Wambach u. a. 2018, S. 27, Fn. 74).

In der Tat empfehlen Dulleck/Kerschbamer/Sutter (2011) in der Gesamtzusammenfassung ihrer Studie, theoretische Modellergebnisse stets verhaltenswissenschaftlich zu prüfen, insbesondere wenn es darum geht, wirtschaftspolitische Maßnahmen abzuleiten und zu empfehlen (S. 553). In einer späteren Arbeit zeigen dieselben Autoren (2017) ergänzend, dass heterogene soziale Präferenzen der Anbieter sowohl die im Vergleich zur Theorie schlechtere Markteffizienz bei Verifiability als auch die bessere bei fehlender Verifiability erklären können. Damit bestätigen sie, dass es sich um verhaltensökonomische Besonderheiten handelt, da die sozialen Präferenzen das Anbieterverhalten bestimmen. Und Dulleck/Kerschbamer/Sutter (2011) räumen dabei zwar ein, dass ihre experimentellen Ergebnisse überraschend seien, wollen dies jedoch keineswegs als Widerspruch zu ihren theoretischen Vorarbeiten in Form des Grundmodells verstanden wissen: Die zunächst unerwarteten Unterschiede in der Wirkung von Liability und Verifiability folgten daraus, dass eine asymmetrische Robustheit der beiden Merkmale in Bezug auf die Koexistenz von Anbietern mit heterogenen Verteilungspräferenzen vorläge (S. 552f.). Im Gegensatz zu Haftungsregeln (Liability), die ihre Wirkung bezüglich Markteffizienz unabhängig von einer möglichen Heterogenität der sozialen Präferenzen der Anbieter entfalteten, wirkten sich im Falle von Nachprüfbarkeit (Verifiability) heterogene soziale Anbieterpräferenzen sowohl bei Fehlen als auch bei Vorliegen von Verifiability im Vergleich zur theoretisch begründbaren Markteffizienz negativ aus.

Kurz gesagt gilt also: Was unten den Annahmen des theoretischen Modells gleich gut wirken müsste, erweist sich in der verhaltenswissenschaftlichen Überprüfung unter der Zusatzannahme heterogener sozialer Anbieterpräferenzen als unterschiedlich wirksam. Es liegt kein Widerspruch vor, im ursprünglichen Modell wird implizit von homogenen sozialen Anbieterpräferenzen ausgegangen, welche sich als unrealistisch erweisen. Da heterogene soziale Anbieterpräferenzen niemals auszuschließen sind, ist Anbieterhaftung der überlegene Mechanismus. Damit liegen klare und konsistente Forschungsergebnisse vor, die durchaus oder sogar gerade wegen ihrer verhaltenswissenschaftlichen Bestätigung bzw. Konkretisierung als Grundlage wirtschaftspolitischer Empfehlungen herangezogen werden können. Haftungsregeln sind ganz klar als besonders wirksam für Effizienzsteigerungen zu empfehlen (vgl. dazu ebenfalls Chen/Li/Zhang 2017).

Verhaltenswissenschaftliche Überprüfungen ökonomischer Theorien insbesondere in Form von Experimenten sind noch eine relativ junge Form der

wissenschaftlichen Arbeitsweise. Sowohl die Verhaltensökonomik (*behavioral economics* (Thaler 2015)) als auch wirtschaftswissenschaftliche Experimente (*experiments in economics* bzw. *experimental economics* (Hey 2018)) wurden erst in den letzten Jahrzehnten verstärkt als ökonomisches Forschungsgebiet bzw. wirtschaftswissenschaftliche Methodik eingesetzt. Seitdem haben sie eine beachtliche Entwicklung genommen und ihre Verbreitung hat sich vervielfacht. Entsprechend mehr Akzeptanz und Anerkennung haben sie daher in den letzten Jahren gewinnen können.

2. Übertragung auf den Markt für Jahresabschlussprüfungen

Die letzte große Reform des europäischen Abschlussprüferrechts war erst 2014. In Deutschland wurde 2015 nach sehr kontroverser Diskussion das Abschlussprüfungsreformgesetz (AReG) und das Abschlussprüferaufsichtsreformgesetz (APAReG) verabschiedet (2016 in Kraft getreten). Erst im Dezember 2018 gab es Schlagzeilen, das Großbritannien den Wirtschaftsprüfern mit harter Regulierung droht, da die Wettbewerbsbehörde CMA und eine vom Londoner Wirtschaftsministerium eingesetzte Kommission zeitgleich empfahlen, zur Kontrolle der Wirtschaftsprüfer – wie in Deutschland die neue Abschlussprüferaufsichtsstelle (APAS) – eine neue Aufsichtsbehörde zu schaffen.

a. Grenzen der Übertragung der generellen Vertrauensgutforschung

Der Markt für Abschlussprüfungen weist eine einzigartige Kombination von Charakteristika auf, die Vergleiche mit anderen Güter- bzw. Dienstleistungsmärkten erschwert: So kommt Abschlussprüfungen eine ganz besondere Rolle zur Sicherung der Transparenz des Kapitalmarkts zu, einen erheblichen Teil stellen Pflichtprüfungen dar und auf Anbieterseite herrscht eine hohe Konzentration, was häufig zu wirtschaftspolitischen Diskussionen bzw. Regulierungsbemühungen führt (Gerakos/Syversen 2015).

Eine Übertragung des Grundmodells von Dulleck/Kerschbamer (2006) auf den Markt für Prüfungsleistungen ist aufgrund dieser Besonderheiten nicht so einfach möglich. Die identifizierten Prämissen bereiten Probleme. Es bestehen Zweifel an der Annahme, dass die eine Jahresabschlussprüfung nachfragenden Unternehmen *homogen* im Sinne des Modells sind, sich also in nur zwei Gruppen (kleines Problem versus großes Problem) einteilen lassen. Auch die Commitment-Bedingung ist nicht erfüllt, da ein Angebot von einer Wirtschaftsprüfungsgesellschaft eingeholt werden kann, ohne dass eine Verpflichtung für das Unternehmen

entsteht, bei Angebotserhalt die Wirtschaftsprüfungsgesellschaft auch tatsächlich zu beauftragen. Verifiability kann eher bejaht werden, da das Unternehmen weitgehend überprüfen kann, welche Leistung es erhalten hat, allerdings nicht so einfach wie im Modell, da es dort ja nur zwei Arten von möglichen Behandlungen gibt (für die angenommenen zwei Gruppen von Nachfragern). Haftungsregeln (Liability) existieren zwar, jedoch nicht in der simplen, im Modell vorgesehenen Form. In Deutschland gilt für gesetzliche Pflichtprüfungen eine gesetzliche Beschränkung der Haftung: § 323 Abs. 2 S. 1 HGB sieht für den Fall fahrlässigen Handelns des Jahresabschlussprüfers eine Ersatzpflicht vor, die pro Prüfung bis zu einer Million Euro betragen kann. Die Geltendmachung von Schadensersatzansprüchen ist damit doppelt begrenzt: Erstens auf Fälle des nachgewiesenen fahrlässigen Handelns des Abschlussprüfers und zweitens in Bezug auf den Betrag.

Selbst wenn das Grundmodell von *Dulleck/Kerschbamer* (2006) nicht einfach auf den Markt für Jahresabschlussprüfungen übertragen werden kann, zeigen die im Modell identifizierten Effizienzbedingungen doch sehr gut, worauf es bei der Milderung der Vertrauensgutprobleme durch Marktmechanismen grundsätzlich ankommt und darüber hinaus, dass Effizienz auf diesen Märkten durch wenige entscheidende Bedingungen erreicht werden kann.

Aus der Modellerweiterung von *Dulleck/Kerschbamer/Sutter* (2011) sind zudem weitere Folgerungen möglich. Die beiden im Modell einflussreichsten Merkmale, Wettbewerb und Haftungsregeln, sind die zwei am intensivsten diskutierten Themen der Wirtschaftsprüfungsbranche. Die fortgesetzte Konzentration auf wenige internationale Wirtschaftsprüfungsgesellschaften (mittlerweile statt Big6 nur noch Big4) löst immer wieder Regulierungsaktivitäten auf den Plan, wie etwa die Prüferrotation, deren wissenschaftliche Rechtfertigung zweifelhaft ist (*Böcking/Gros* 2015; *Gerakos/Syverson* 2015; *Hennrichs* 2015). Und über die Haftungsregeln wird immer wieder stark diskutiert. In Deutschland gibt es die limitierte Ersatzpflicht bei Fahrlässigkeit (*Doralt* 2015, S. 292 ff.), wobei die Begrenzung der Haftung insbesondere in den Fällen grob fahrlässigen Handelns des Wirtschaftsprüfers seit Jahren kritisiert wird (vgl. *Kilian/Rimkus* 2016)²².

Dulleck/Kerschbamer/Sutter (2011) ziehen Schlüsse auf die Gründe, warum Haftungsregeln (Liability) auf Vertrauensgutmärkten besser wirken als Nachprüf-

²² Im Fokus dieser Auseinandersetzung um die Haftungsbegrenzung steht § 54a Abs. 1 Nr. 2 WPO, der die vertragliche Begrenzung von Ansprüchen auf Ersatz eines fahrlässig verursachten Schadens durch vorformulierte Vertragsbedingungen (oft in Form der Allgemeinen Auftragsbedingungen (AAB)) regelt. Gegen die Gewährung einer Haftungsprivilegierung in Fällen der groben Fahrlässigkeit spricht sich eine erhebliche Zahl von Stimmen aus. Zu den allgemeinen Auftragsbedingungen und deren Anpassung vgl. *Ritzer-Angerer* (2017).

barkeit (Verifiability), und argumentieren, dass Haftung den Kunden wirksam vor Undertreatment schützt (s. o.). Verifiability hingegen kann nur wirksam vor Overcharging schützen und Overtreatment spielt – zumindest bei geringfügigem Umfang – kaum eine Rolle, da der Kunde keine Information darüber erhält, dass unnötige Eingriffe vorgenommen wurden, so dass er ungebrochen zufrieden ist. Dies lässt sich eindeutig auf die Jahresabschlussprüfung übertragen, da auch hier nur Undertreatment-Fälle wirklich problematisch sind, wenn der Prüfer also Prüfungsschritte nicht hinreichend ausführlich oder sorgfältig vorgenommen hat, Fehler oder Unstimmigkeiten übersehen wurden. Wird hingegen ein Prüfungsschritt doppelt gemacht und abgerechnet oder ggf. zu viele Stunden abgerechnet, entstehen dem Kunden zwar höhere Kosten, aber wiederum, besonders wenn dies nur in geringem Umfang vorkommt, fällt dies weder auf noch besonders ins Gewicht.

Undertreatment hingegen kann bei Jahresabschlussprüfungen (sog. Under-Auditing) als natürliche Folge der Vertrauensguteigenschaften von Prüfungsleistungen verstanden werden, da der Vorteil der Kostenersparnis durch verringerte Prüfungsanstrengungen dem weniger schwerwiegenden, weil unwahrscheinlichen Nachteil gegenübersteht, dass das Under-Auditing tatsächlich entdeckt wird (*Causholli/Knechel/Lin/Sappington* 2013, S. 587 u. Endnote 36).

b. Einordnung der Vertrauensgutforschung in die wirtschaftswissenschaftliche Literatur zur Abschlussprüferhaftung

Die Literatur zu Vertrauensgütern ist erst wenige Jahre alt, (die neusten, teilweise experimentellen Arbeiten erst ein oder zwei Jahre) und zur Anwendung auf die Jahresabschlussprüfung sind bislang vor allem die Arbeiten von *Causholli* et al. vorhanden, die sich ausschließlich mit den Vertrauensguteigenschaften der Jahresabschlussprüfung befasst und nicht mit der Frage, was das für die eventuell zu empfehlende Marktregulierung bedeuten könnte. Die konkrete Frage, ob sich mit der in der Vertrauensgutliteratur gezeigten Dominanz der Anbieterhaftung eine neue Regulierung der Wirtschaftsprüferhaftung begründen lässt, ist meines Wissens bisher kaum wissenschaftlich bearbeitet worden. Hier besteht zukünftiger Forschungsbedarf.

Die wissenschaftliche Literatur zu der Frage, wie Wirtschaftsprüferhaftung grundsätzlich auszugestalten sei, ist uneinheitlich, zuweilen sogar widersprüchlich. Dies liegt an der Vielfalt der Modelle, die angewendet werden (können) und daran, dass sich oft zeigt, dass die Ergebnisse sehr sensitiv auf Modellprämissen reagieren (*Koch/Schunk* 2013). *Schäfer* hat bezüglich der Fehlnutzung von Informationen und der Haftung für fehlerhafte Wertgutachten interessante Über-

legungen veröffentlicht und dabei insbesondere die Relevanz der Unterscheidung von Ressourcen- und reinen Umverteilungsschäden herausgearbeitet (*Schäfers* 2002). Spätere Arbeiten mit *Bigus* nehmen Bezug auf die Haftung des Wirtschaftsprüfers (*Bigus/Schäfer* 2007; dazu ebenso *Ott* 2008). Neuere Arbeiten verteidigen Haftungsbeschränkungen verhaltenswissenschaftlich (z. B. wieder *Bigus* 2015 und 2016) oder mit dem Argument günstigerer Preise für die Prüfung (so etwa *Louis u. a.* 2019). Günstige Preise bedeuten höherer Wettbewerbsdruck, worauf die Prüfer „mit einer weiteren Erhöhung der Effizienz ihrer Prüfungsabwicklung“ (*APAK* 2012) reagieren, was wiederum tendenziell zu unsorgfältiger Prüfungsdurchführung führt, die wiederum das Auftreten einfach fahrlässiger bzw. leichtfertiger Prüfungsfehler begünstigt, usw. Diese Spirale trägt einen Anteil zur Erklärung der Bilanzskandale²³ bei, was das grundsätzlich gerechtfertigte ökonomische Argument der günstigen Preise in Zweifel zieht, (vgl. dazu *Bauer* 2017). Stattdessen wäre vor diesem Hintergrund eine Diskussion darüber gerechtfertigt, welchen Wert die Gesellschaft einer höheren Prüfungsqualität beimisst bzw. beimessen will, durchaus in Konkurrenz zur Finanzierung anderer öffentlicher Güter, womit man die Bezahlung der Honorare durch das beauftragende Unternehmen auf den Prüfstand hebt. Es bleibt unverständlich, warum sich der Berufsstand selbst sich diese Argumentation nicht zu eigen macht: Qualität hat ihren Preis, auch in der Abschlussprüfung. Neue Dissertationen stellen Überlegungen zur Straffälligkeit und Strafbedürftigkeit von beruflichem Fehlverhalten des Abschlussprüfers in den Raum (*Prinz* 2013; *Bauer* 2017), welche eine neue Strenge mit dem Berufsstand befördern, die sich natürlich auch auf Haftungsfragen auswirken wird.

Erschwert wird die wissenschaftliche Diskussion grundsätzlich durch das enorme Ausmaß der Lobbyarbeit des internationalen Berufsstandes (vgl. dazu *Allen/Ramanna/Roychowdhury* 2018). Ein erheblicher Teil des Prüfungsschriftums wurde von Berufspraktikern verfasst und der Berufsstand macht erhebliches Interesse an einer Haftungsbegrenzung geltend,²⁴ wie bereits die EU-Reform der Wirtschaftsprüferhaftung 2008 gezeigt hat. Begründet wurde diese Reform mit der

23 Ein Bilanzskandal folgt aus aufgeklärtem unethischen Verhalten von Menschen innerhalb oder im Auftrag von Unternehmen. Wird der Jahresabschluss durch bewusste Manipulationen geändert oder sogar absichtlich gefälscht, liegt Bilanzbetrug vor. Bilanzbetrug gibt es solange wie Bilanzen selbst. Dennoch gab es im neuen Jahrhundert aufsehenerregende Bilanzskandale durch Bilanzfälschungen wie Enron in USA oder FlowTex in Deutschland (vgl. *Peemöller et al.* 2016; *Müller* 2004). Ziel einer Bilanzmanipulation ist es, den Kapitalmarkt (aktuelle und potentielle Anleger) oder andere Adressaten zu täuschen.

24 Dies gilt in besonderen Maße für die großen internationalen Prüfungsgesellschaften, was auch darin begründet ist, dass sich für sie unbegrenzte Haftung besonders intensiv auswirken würde, vgl. *Johnson/Stokes/Watts* (1995).

Situation des Oligopolmarktes unter den großen internationalen Abschlussprüfungsgesellschaften und den dramatischen Konsequenzen des Ausscheidens einer weiteren von diesen sowie mit vorgebrachten enormen Haftungsrisiken, die das Überleben derselben Prüfungsgesellschaften bedrohten. *Hopt* kritisiert den damaligen EU-Kommissar für Aktien- und Wirtschaftsprüferrecht *McCreevey*, der sein Amt ausschließlich für die Installation dieser Schutzregelung für Wirtschaftsprüfer nutzte und selbst als Wirtschaftsprüfer tätig ist (*Hopt* 2015, S. 198). In Deutschland hatte die Reform von 2008 nahezu keine Auswirkung, da die Regelungen zur Haftungsbegrenzung die neuen EU-Vorgaben bereits erfüllten und somit einfach beibehalten wurden. In den anderen europäischen Ländern waren die Auswirkungen der Reform 2008 unterschiedlich (*Samsonova-Taddei/Humphrey* 2015). Vor diesem Hintergrund haben sich die Reformvorschläge von 2011 erst gar nicht auf das Terrain des Haftungsrechts gewagt (*Hopt* 2015, S. 198) und dennoch war der Widerstand des Berufsstands und die entfachten Diskussionen erheblich, wenn freilich – zumindest in Bezug auf die tatsächlichen Mängel der Reform – berechtigt.

c. Empirische Belege für die Übertragbarkeit der Vertrauensgutforschung

Empirische Studien für Wirkungen von unterschiedlichen Haftungsregimen auf die Prüfungsqualität sind rar. Hier besteht ebenfalls noch sehr großer Forschungsbedarf. In einer ganz neuen Studie untersuchen *Louis u. a.* (2019) Prüfungsfälle von PWC – zum einen Fälle ohne Haftungsbeschränkungsklauseln (bei öffentlich gehandelten Unternehmen) und zum anderen Fälle mit solchen Klauseln – und kommen zu dem Ergebnis, dass Haftungsbeschränkungsklauseln Prüferhonorare reduzierten, sich dies aber so gut wie gar nicht auf die Prüfungsqualität auswirke. Damit liegt ein erster empirischer Hinweis zur Wirkung von Haftungsbeschränkungen vor, für eine gesicherte empirische Evidenz ist dies jedoch bei weitem noch nicht ausreichend. Es braucht empirische Belege, die zusätzlich andere Wirtschaftsprüfungsgesellschaften miteinbeziehen, ganz unterschiedliche geprüfte Unternehmen und zudem sind Unterschiede zwischen vertraglicher und gesetzlicher Haftungsbeschränkung bezüglich ihrer Wirkung auf den Markt zu erwarten und daher zu berücksichtigen.

3. Offene Fragen zur Wirtschaftsprüferhaftung

Neben der zentralen Frage, ob Wirtschaftsprüferhaftung begrenzt oder unbegrenzt gestaltet werden sollte, gibt es zahlreiche weitere Fragestellungen, die eine bedeutende Rolle in diesem Kontext spielen. Zentrale offene Fragen sollen

in diesem Beitrag an dieser Stelle ohne Anspruch auf Vollständigkeit genannt werden, nicht zuletzt um darauf hinzuweisen, dass die wissenschaftliche Frage nach der optimalen Ausgestaltung der Wirtschaftsprüferhaftung von solchen Nebenkriegsschauplätzen zum Teil abhängt und erschwert wird.²⁵

- Welche Rolle spielt die Gestaltung der Haftung der anderen für die Richtigkeit des Jahresabschlusses Verantwortlichen (bei AG Vorstand und Aufsichtsrat) für die Gestaltung der Wirtschaftsprüferhaftung? (*Hennrichs* 2015, S. 249);
- Hat die vertragliche Haftungsvereinbarung (oder Vereinbarung über AGB bzw. AAB) des Wirtschaftsprüfers mit seinem Klienten Vor- oder Nachteile gegenüber einer gesetzlichen Regelung? (*Goette* 2015, S. 322, 324);
- Welchen Einfluss hat die konkrete Höhe der gesetzlich begrenzten Haftungssumme bei begrenzter Haftungsübernahme? (*Hopt* 2015, S. 198; *Goette* 2015, S. 323);
- Ist das Beibehalten einer starren Haftungsbegrenzung bei evtl. zu geringen Haftungssummen mit Nachteilen verbunden wie die Etablierung von Surrogaten, die die Lücke zu schließen versuchen? (*Hopt* 2015, S. 198);
- Ist es sinnvoll, die stark unterschiedlichen europäischen Haftungsregime des Wirtschaftsprüfers zu harmonisieren oder sollte vielmehr überlegt und diskutiert werden, ob das jeweilige Privatrecht zum Schutz der Anleger eingesetzt werden sollte? (*Goette* 2015, S. 322 f.);
- Ist die Politik der Bundesrepublik, sich der Harmonisierung des europäischen Haftungsrechts zu lange zu verweigern, kurzfristig? (*Hopt* 2015, S. 200).

V. Resümee und Ausblick: Bedeutung der Vertrauensgutforschung für die Wirtschaftsprüferhaftung

Die Jahresabschlussprüfung weist nicht nur Vertrauensguteigenschaften auf, sie zählt zu den typischen Vertrauensgütern, da die Jahresabschlussprüfung ausschließlich von sog. Expert providers angeboten wird. Für Vertrauensgüter konnte

²⁵ Die hier zusammengetragenen Fragen entstammen dem Dialog, der im Rahmen eines Sondersymposiums im November 2013 stattfand, welches die Herausgeber der Zeitschrift für Unternehmens- und Gesellschaftsrecht anlässlich der Reform des europäischen Abschlussprüfungsrechts veranstalteten. Dazu wurden Vertreter der Rechtswissenschaft, der Wirtschaftswissenschaften und der Praxis eingeladen. Bedauerlicherweise fehlten Vertreter der Volkswirtschaftslehre, insbesondere der Wirtschaftspolitik, die vorhandene Erkenntnisse zur Wirkung verschiedener Regulierungsoptionen auf den Markt hätten beitragen können.

theoretisch im Modell und verhaltenswissenschaftlich experimentell die Überlegenheit von Anbieterhaftungsregeln in Bezug auf das Ziel der Markteffizienz belegt werden, da sich deren Wirkung – im Gegensatz zur untersuchten Alternative – unabhängig von heterogenen sozialen Anbieterpräferenzen entfaltet. Im Vergleich zu anderen Vertrauensgütern liegt bei der Abschlussprüfung sogar eine verschärfte Vertrauensgutproblematik vor: Der Abschlussprüfer hat nicht nur die für Vertrauensgüter typischen Optionen, seinen Informationsvorsprung opportunistisch auszunutzen, er arbeitet bei der Erbringung seiner Dienstleistung eng mit den Managern zusammen, die aufgrund ihrer Prinzipal-Agent-Beziehung zu den Eigentümern/Kunden Anreize haben, ihre Erfolge herauszustellen und ihre Misserfolge zu verbergen. Es sind wissenschaftliche Belege dafür vorhanden, dass der Abschlussprüfer das in ihn von der Öffentlichkeit gesetzte Vertrauen deshalb mit deutlich höherer Wahrscheinlichkeit verletzen wird als seine Beziehung zum Management (vgl. *Neu* 1991).

Dieser Beitrag wendet erstmals die Erkenntnisse der Vertrauensgutforschung auf die Jahresabschlussprüfung und die Wirtschaftsprüferhaftung an. Dabei zeigt sich, dass vor ca. zehn Jahren in Europa eine unterschiedlich ausgestaltbare Haftungsbegrenzung für Wirtschaftsprüfer verabschiedet wurde und dass sich die jüngste EU-Reform des Abschlussprüferrechts nicht auf das Terrain des Haftungsrechts begeben hat. Der Kritik an Details der Haftungsbegrenzung insbesondere von juristischer Seite stehen eindeutige Interessen an deren Verteidigung durch die Berufsverbände gegenüber. Die Vertrauensgutforschung kann in diesem Kontext als wissenschaftlich begründetes Plädoyer für eine strenge Anbieterhaftung dienen. Der vielfach vorgetragene Einwand, auf den Prüfungsmarkt sei besondere Vorsicht geboten, da er viele Besonderheiten aufweise und vor allem bereits sehr stark reglementiert sei, kann dabei durchaus Rechnung getragen werden.²⁶ Die empirische Forschung reicht bei weitem noch nicht aus, um für den Prüfungsmarkt konkrete Empfehlungen abzugeben. Hier besteht noch großer Forschungsbedarf.

Trotz der erst wenige Jahre zurückliegenden europäischen Reform des Abschlussprüfungsrechts und des Abschlussprüferaufsichtsrechts bestehen nach wie vor viele offene Fragen in Bezug auf die von Europa angestrebte Harmonisierung der Haftungsregime der Mitgliedsstaaten. Dabei ist bereits im Vorfeld klar, dass eine Harmonisierung nicht alle Fragen lösen können wird. Dennoch besteht in der EU ein einheitlicher Binnenmarkt sowie ein einheitlicher Kapital-

²⁶ Auf stark reglementierten Märkten ist das Betrugspotential zunächst grundsätzlich geringer. Second-Degree-Moral-Hazard-Effekte werden jedoch gerade auf stark reglementierten Märkten (zum Beispiel dem Taximarkt) analysiert. Es besteht trotz starker Reglementierung die Gefahr von Overtreatment und Overcharging (vgl. *Kerschbamer/Neururer/Sutter* (2016), S. 7455).

markt, so dass die Diskussion um die Wirtschaftsprüferhaftung nicht divers geführt werden sollte. Und selbst wenn sich Europa in absehbarer Zeit zu einer einheitlichen Regulierung in Bezug auf die Wirtschaftsprüferhaftung durchringen sollte, warten zukünftig noch weitere Aufgaben, so etwa die Internationalisierung der Prüfungsstandards. Dafür scheint die Zeit noch nicht reif wie schon die Divergenz zwischen den europäischen Mitgliedsstaaten zeigt.

Literatur

- Akerlof, George A. (1970), „The market for “lemons”: Quality uncertainty and the market mechanism”, *Quarterly Journal of Economics*, Jg. 84, H. 3, S. 488–500.
- Allen, Abigail M., Karthik Ramanna und Sugata Roychowdhury (2018), „Auditor lobbying on accounting standards”, *Journal of Law, Finance & Accounting*, Jg. 3, H. 2, S. 291–331.
- Abschlussprüferaufsichtskommission (APAK), „Qualität der Abschlussprüfung – Tätigkeitsbericht 2012 der APAK“, *WP Praxis* 7/2013, S. 135–136.
- Baetge, Jörg und Sonja (2003), „Normative Überlegungen zur Ausgestaltung der Abschlussprüfung im Lichte aktueller Unternehmenskrisen“, in: Peter Wollmert, Norbert Schönbrunn, Udo Jung, Hilmar Siebert und Michael Henke (Hrsg.), *Wirtschaftsprüfung und Unternehmensüberwachung: Festschrift für Prof. Dr. Dr. h.c. Wolfgang Lück*, Düsseldorf, S. 179–203.
- Ballwieser, Wolfgang (2008), „Entwicklung und Problemfelder von Wirtschaftsprüfungsgesellschaften“, in: Wolfgang Ballwieser und Wolfgang Grewe (Hrsg.), *Wirtschaftsprüfung im Wandel, Herausforderungen an Wirtschaftsprüfung, Steuerberatung, Consulting und Corporate Finance, Festgabe 100 Jahre Südtreu/Deloitte*, München, S. 1–16.
- Balafoutas, Loukas, Rudolf Kerschbamer und Matthias Sutter (2017), „Second-degree moral hazard in a real-world credence goods market”, *The Economic Journal*, Jg. 127, H. 599, S. 1–18.
- Bauer, Thea C. (2017), „Die Neuregelung der Strafbarkeit des Jahresabschlussprüfers, Ein Vorschlag de lege ferenda zur Erfassung der straffälligen und strafbedürftigen Konstellationen von beruflichem Fehlverhalten des Jahresabschlussprüfers“, Tübingen.
- Berg, Hartmut, Dieter Cassel und Karl-Hans Hartwig (2007), *N. Theorie der Wirtschaftspolitik*, in: *Vahlens Compendium der Wirtschaftstheorie und Wirtschaftspolitik*, Bd. 2, München Vahlen.
- Bester, Helmut und Matthias Dahm (2017), „Credence goods, costly diagnosis and subjective evaluation”, *The Economic Journal*, Jg. 128, H. 611(Juni), S. 1367–94.
- Bester, Helmut und Yaofu Ouyang (2018), *Optimal procurement of a credence good under limited liability*, *Diskussionsbeiträge*, No. 2018/1, online verfügbar unter <https://www.econstor.eu/handle/10419/173277>, zuletzt abgerufen am 26.6.2020.
- Bigus, Jochen (2015), „Loss aversion, audit risk judgments, and auditor liability”, *European Accounting Review*, Jg. 24, H. 3, S. 581–606.
- Bigus, Jochen (2016), „Optimism and auditor liability”, *Accounting and Business Research*, Jg. 46, H. 6, S. 577–600.

- Bigus, Jochen und Hans-Bernd Schäfer (2007), „Die Haftung des Wirtschaftsprüfers am Primär- und am Sekundärmarkt – eine rechtsökonomische Analyse“, *Zeitschrift für Betriebswirtschaft*, Jg. 77, H. 1, S. 19–49.
- Birchler, Urs und Monika Bütler (2007), *Information economics*. London: Routledge.
- Bluethgen, Ralph, Steffen Meyer und Andreas Hackethal (2008), *High-Quality Financial Advice Wanted!*, Working Paper, online verfügbar unter <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.596.2310&rep=rep1&type=pdf>, zuletzt abgerufen am 26.6.2020.
- Böcking, Hans-Joachim und Marius Gros (2015), „Reform des Abschlussprüfermarkts – Erkenntnisse der Betriebswirtschaftslehre und evidenzbasierte Regulierung“, *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, Jg. 44, H. 2, S. 305–19.
- Bonroy, Olivier und Christos Constantatos (2008), „On the use of labels in credence goods markets“, *Journal of Regulatory Economics*, Jg. 33, H. 3, S. 237–52.
- Causholli, Monika (2009), *Audits as credence goods: What do auditors know and how do they use their information?*, Diss., University of Florida.
- Causholli, Monika und W. Robert Knechel (2009), *Is the audit a credence service?*, Working Paper, University of Florida.
- Causholli, Monika, W. Robert Knechel, Haijin Lin und David E.M. Sappington (2013), „Competitive procurement of a credence good: The case of auditing“, *European Accounting Review*, Jg. 22, H. 3, S. 573–605.
- Causholli, Monika und W. Robert Knechel (2012), „An Examination of the Credence Attributes of an Audit“, *Accounting Horizons*, Jg. 26, H. 4, S. 631–56.
- Chen, Yongmin, Jianpei Li und Jin Zhang (2017), *Liability in markets for credence goods*. Munich Personal RePEc Archive Paper No. 80206, online verfügbar unter https://mpra.ub.uni-muenchen.de/80206/1/MPRA_paper_80206.pdf, zuletzt abgerufen am 26.6.2020.
- Chiappori, Pierre-André, Franck Durand und Pierre-Yves Geoffard (1998), „Moral hazard and the demand for physician services: first lessons from a French natural experiment“, *European Economic Review*, Jg. 42, H. 3–5, S. 499–511.
- Chiappori, Pierre-André und Bernard Salanié (2000), „Testing for asymmetric information in insurance markets“, *The Journal of Political Economy*, Jg. 108, H. 1, S. 56–78.
- Canadian Institute of Chartered Accountants (CICA), *Report of the Commission to study the public's expectation of audits*, June 1988, online verfügbar unter <https://wamacdonald.com/wp-content/uploads/2014/07/Macdonald-Commission-Report-June-1988.pdf>, zuletzt abgerufen am 26.6.2020.
- Coenenberg, Adolf, Axel Haller und Wolfgang Schultze (2018), *Jahresabschluss und Jahresabschlussanalyse*, 25. Aufl., Stuttgart, Schäffer-Poeschel.
- Darby, Michael R. und Edi Karni (1973), „Free competition and the optimal amount of fraud“, *Journal of Law and Economics*, Jg. 16, H. 1, S. 67–81.
- Doralt, Werner (2015), „Die Haftung des gesetzlichen Abschlussprüfers – Mitverschulden, Ansprüche Dritter und Wege der Haftungsbegrenzung“, *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, Jg. 44, H. 2, S. 266–304.
- Doran, Evan, Jane Robertson und David Henry (2005), „Moral hazard and prescription medicine use in Australia – the patient perspective“, *Social Science & Medicine*, Jg. 60, H. 7, S. 1437–43.
- Dulleck, Uwe und Rudolf Kerschbamer (2006), „On Doctors, Mechanics, and Computer Specialists: The Economics of Credence Goods“, *Journal of Economic Literature*, Jg. 44, H. 1, S. 5–42.

- Dulleck, Uwe, Rudolf Kerschbamer und Matthias Sutter (2011), „The Economics of Credence Goods: An Experiment on the Role of Liability, Verifiability, Reputation, and Competition”, *American Economic Review*, Jg. 101, H. 2, S. 526–55.
- Drukarczyk, Jochen (1993), *Theorie und Politik der Finanzierung*, 2. Aufl., München, Vahlen.
- Einav, Liran, Amy Finkelstein, Stephen R. Ryan, Paul Schrimpf und Mark R. Cullen (2013), „Selection on moral hazard in health insurance”, *American Economic Review*, Jg. 103, H. 1, S. 178–219.
- Eisenhardt, Patrick (2011), *Überwachung der Abschlussprüfer*, Hohenheim.
- Emons, Winand (1997), „Credence Goods and Fraudulent Experts”, *RAND Journal of Economics*, Jg. 28, H. 1, S. 107–19.
- Fleischer, Holger (2002), „Gutachten F zum 64. Deutschen Juristentag“, in: Holger Fleischer und Hanno Merkt (Hrsg.), *Gutachten F + G zum 64. DJT*, München, S. F 1–F 142.
- Fritsch, Michael (2011), *Marktversagen und Wirtschaftspolitik*, 8. Aufl., München, Vahlen.
- Gerakos, Joseph J. und Chad Syverson (2015), „Competition in the Audit Market: Policy Implications”, *Journal of Accounting Research*, Jg. 53, H. 4, S. 725–75.
- Goette, Martin (2015), „Bericht über die Diskussion“, *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, Jg. 44, H. 2, S. 320–24.
- Gottschalk, Felix C.H. (2017), *Essays on credence goods markets with applications to health care and financial advice*, ETH Zürich Library, online verfügbar unter <https://www.research-collection.ethz.ch/bitstream/handle/20.500.11850/235915/Gottschalk%20-%20Essays%20on%20Credence%20Goods%20Markets%20-%202017.pdf?sequence=1&isAllowed=y>, zuletzt abgerufen am 26.6.2020.
- Gottschalk, F.C.H. (2018), *What characterizes credence goods? A critical look at the literature*. Working Paper, online verfügbar unter https://www.researchgate.net/publication/323466297_What_Characterizes_Credence_Goods_A_Critical_Look_at_the_Literature, zuletzt abgerufen am 26.6.2020.
- Hadfield, Gillian K. (2000), „The price of law: How the market for lawyers distorts the justice system”, *Michigan Law Review*, Jg. 98, H. 4, S. 953–1006.
- Henrichs, Joachim (2015), „Abschlussprüferreform im Unionsrecht“, *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, Jg. 44, H. 2, S. 248–65.
- Herkendell, Anja (2007), „Regulierung der Abschlussprüfung: Eine Wirksamkeitsanalyse zur Wiedergewinnung des öffentlichen Vertrauens“, Wiesbaden.
- Hey, John D. (2018), *Experiments in Economics: Decision Making and Markets*, New Jersey.
- Hilger, Nathaniel G. (2016), „Why Don't People Trust Experts?“, *Journal of Law and Economics*, Jg. 59, H. 2, S. 293–311.
- Hopt, Klaus J. (2015), „Abschlussprüfung in Deutschland und Europa nach der europäischen Reform von 2014“, *Zeitschrift für Unternehmens- und Gesellschaftsrecht*, Jg. 44, H. 2, S. 186–203.
- Johnson, Ronald B., Donald J. Stokes und David G. Watts (1995), „Auditor preferences for liability limitation”, *Accounting & Finance*, Jg. 35, H. 2, S. 135–153.
- Kerschbamer, Rudolf, Daniel Neururer und Matthias Sutter (2016), „Insurance coverage of customers induces dishonesty of sellers in markets for credence goods”, *PNAS* 113(27), S. 7454–58.
- Kerschbamer, Rudolf und Matthias Sutter (2017), „The economics of credence goods – a survey of recent lab and field experiments”, *CESifo Economic Studies*, S. 1–23, online verfügbar unter <https://academic.oup.com/cesifo/article/63/1/1/2992734>, zuletzt abgerufen am 26.6.2020.

- Kerschbamer, Rudolf, Matthias Sutter und Uwe Dulleck (2017), „How social preferences shape incentives in (experimental) markets for credence goods“, *The Economic Journal*, Jg. 127, H. 600, S. 393–416.
- Kilian, Matthias und Felix Rimkus (2016), „Die Verwendung Allgemeiner Auftragsbedingungen für Wirtschaftsprüfer und Wirtschaftsprüfungsgesellschaften – Klauselrechtliche Probleme in der Praxis“, *Zeitschrift für Wirtschaftsrecht*, H. 13, S. 608–13.
- Knechel, W. Robert und Marleen Willekens (2006), „The role of risk management and governance in determining audit demand“, *Journal of Business Finance & Accounting*, Jg. 33, H. 9–10, S. 1344–67.
- Knechel, W. Robert, Paul Rouse und Caren Schelleman (2009), „A modified audit production framework: Evaluating the relative efficiency of audit engagements“, *The Accounting Review*, Jg. 84, H. 5, S. 1607–38.
- Koch, Christopher und Daniel Schunk (2013), „Limiting liability? – Risk and ambiguity attitudes under real losses“, *Schmalenbach Business Review*, Jg. 65, H. 1, S. 54–75.
- Kremer, René M. (2007), „Die zivilrechtliche Verantwortlichkeit des Wirtschaftsprüfers gegenüber Kapitalanlegern“, Göttingen.
- Li, Jianpei und Yaofu Ouyang (2016), Expert costs and the role of verifiability. Munich Personal RePEc Archive Paper No. 74390, online verfügbar unter https://mpra.ub.uni-muenchen.de/74390/1/MPRA_paper_74390.pdf, zuletzt abgerufen am 26.6.2020.
- Louis, Henock, Tom Pearson, Dahlia Robinson, Michael Robinson und Amy X. Sun (2019), „The effects of the extant clauses limiting auditor liability on audit fees and overall reporting quality“, *Journal of Empirical Legal Studies*, Jg. 16, H. 2, S. 381–410.
- Lu, Fangwen (2014), „Insurance coverage and agency problems in doctor prescriptions: Evidence from a field experiment in China“, *Journal of Development Economics*, Jg. 106, H. C, S. 156–67.
- Marten, Kai-Uwe, Reiner Quick und Klaus Ruhnke (2015), *Wirtschaftsprüfung*, 5. Aufl., Stuttgart, Schäffer-Poeschel.
- Müller, Christian (2004), „Bilanzskandale – Eine institutionenökonomische Analyse“, *Perspektiven der Wirtschaftspolitik*, Jg. 5, H. 2, S. 211–226.
- Nelson, Phillip (1970), „Information and Consumer Behavior“, *Journal of Political Economy*, Jg. 78, H. 2, S. 311–29.
- Neu, Dean (1991), „Trust, impression management and the public accounting profession“, *Critical Perspectives on Accounting*, Jg. 3, H. 2, 295–313.
- O’Keefe, Terrence B., Dan A. Simunic und Michael T. Stein (1994), „The production of audit services – Evidence from a major public accounting firm“, *Journal of Accounting Research*, Jg. 32, H. 2, S. 241–61.
- OLG Düsseldorf, Urteil vom 19. November 1998, 8 U 59/98, juris.
- Ott, Claus (2008), „Die Haftung von Wirtschaftsprüfern gegenüber Anlegern am Kapitalmarkt“, in: Thomas Eger, Jochen Bigus, Claus Ott und Georg von Wangenheim (Hrsg.), *Internationalisierung des Rechts und seine ökonomische Analyse: Festschrift für Hans-Bernd Schäfer zum 65. Geburtstag*, Wiesbaden, S. 171–191.
- Peemöller, Volker H., Harald Krehl, Stefan Hofmann und Jana Lack (2016), *Bilanzskandale: Delikte und Gegenmaßnahmen*, 2. Aufl., Berlin.
- Prinz, Christian (2013), *Grundfragen der Strafbarkeit der Abschlussprüfer bei der Jahresabschlussprüfung einer Kapitalgesellschaft*, Hamburg.
- Ritzer-Angerer, Petra (2017), „Allgemeine Auftragsbedingungen (AAB) für Wirtschaftsprüfer“, *Wirtschaftsprüfung*, Jg. 70, H. 24, S. 1431–36.

- Samsonova-Taddei, Anna und Christopher Humphrey (2015), „Risk and the construction of a European audit policy agenda: The case of auditor liability”, *Accounting, Organizations and Society*, Jg. 41, S. 55–72.
- Sattler, Matthias (2011), *Vereinbarkeit von Abschlussprüfung und Beratung*, Wiesbaden.
- Schäfer, Hans-Bernd (2002), „Haftung für fehlerhafte Wertgutachten in wirtschaftswissenschaftlicher Perspektive“, *Archiv für die civilistische Praxis (AcP)*, Bd. 202, H. 4/5, S. 808–840.
- Schneider, Henry S. (2012), „Agency Problems and Reputation in Expert Services: Evidence from Auto Repair”, *Journal of Industrial Economics*, Jg. 60, H. 3, S. 406–33.
- Stiglitz, Joseph E. (2002), „Information and the Change in the Paradigm in Economics”, *American Economic Review*, Jg. 92, H. 3, S. 460–501.
- Thaler, Richard H. (2015), *Misbehaving: The Making of Behavioral Economics*, New York.
- Wambach, Achim, Dagmar Kollmann, Jürgen Kühling, Thomas Nöcker und Angelika Westerwelle (2018), *Zweiundzwanzigstes Hauptgutachten der Monopolkommission Wettbewerb 2018, Reformen im Vergütungssystem für die Versorgung mit Arzneimitteln durch Großhändler und Apotheke*, in: *Bundestags-Drucks. 19/3300 v. 10.7.2018*, S. 6–37.
- Wiemann, Daniela (2011), *Prüfungsqualität des Abschlussprüfers*, Wiesbaden.
- Wild, Andreas und Ellen-Katharina Scheithauer (2012), „Die Entwicklung der Konzentration auf dem Markt für Abschlussprüfungen unter Berücksichtigung externer Einflüsse“, *Wirtschaftsprüfung*, Jg. 65, H. 4, S. 186–197.
- Wolinsky, Asher (1993), „Competition in a Market for Informed Experts’ Services”, *RAND Journal of Economics*, Jg. 24, H. 3, S. 380–98.
- Zimmer, Daniel, Dagmar Kollmann, Thomas Nöcker, Achim Wambach und Angelika Westerwelle (2014), *Zwanzigstes Hauptgutachten der Monopolkommission 2012/2013, Wettbewerbsdefizite auf Taximärkten*, in: *Bundestags-Drucks. 18/2150 v. 17.7.2014*, S. 114–28.