

05.25

ZIR

60. Jahrgang
Oktober 2025
Seiten 217–260

www.ZIRdigital.de

Zeitschrift Interne Revision

Herausgeber:

DIIR

Deutsches Institut für
Interne Revision e.V.

Fachzeitschrift für Wissenschaft und Praxis

Standards · Regeln · Berufsstand

Global Internal Audit Standards, Standard 14.5:
Gesamturteil zu einem Auftrag 220

*DIIR-Arbeitskreis „Interne Revision in
Öffentlichen Institutionen“*

Künstliche Intelligenz in der Internen Revision 224

Prof. Dr. Katharina Dillkötter · Moritz Hämel

Management · Best Practice · Arbeitshilfen

Künstliche Intelligenz in der Internen Revision:
Eine strategische Notwendigkeit 234

Ekaterina Steklyannikova

Digitalisierung der Internen Revision:
Dokumentenanalyse mit KI 240

Dr. Tobias Schnier

Wissenschaft · Forschung

Data Analytics und Kennzahlen in der Rechnungslegung 246

Prof. Dr. Dr. Joachim Häcker · Pascal Bohn

PROF. DR. DR. JOACHIM HÄCKER · PASCAL BOHN

Data Analytics und Kennzahlen in der Rechnungslegung

Wie der Bilanzbetrug von Wirecard schon früher hätte entdeckt werden können



Prof. Dr. Dr. Joachim Häcker ist Professor an der HM Business School der Hochschule München (HM). Ferner ist er Gründer und Direktor des European Institute of Quantitative Finance (EIQF).

Pascal Bohn ist Tax Associate bei der WERTE + TEAM Steuerberatungsgesellschaft AG und hat an der HM Business School den Master in Finance and Accounting abgeschlossen.

Die frühzeitige Identifikation von Bilanzbetrug stellt eine zentrale Herausforderung für Unternehmen, Investoren und Regulierungsbehörden dar. Auf der einen Seite können Finanzkennzahlen errechnet und miteinander kombiniert werden. Vorreiter ist hier der sogenannte Beneish-M-Score. Auf der anderen Seite kann mittels Data Analytics die Zahlenfolge der Bilanz- und GuV-Zahlen analysiert werden. Der erste Ansatz zeigt im Fall Wirecard keine Anomalien auf. Mit dem zweiten Ansatz zeigt sich, dass der Bilanzbetrug von Wirecard schon viel früher hätte entdeckt werden können.

1. Einleitung

Der größte Steuerbetrug in der Geschichte der Bundesrepublik ist der Cum-Ex-Skandal, welcher den deutschen Steuerzahler circa 32 Milliarden Euro kostete. Der zweitgrößte Betrugsfall und zugleich der größte und wohl auch spektakulärste Bilanzskandal der deutschen Wirtschaftsgeschichte ist der Wirecard-Skandal. Im August 2018 wurde Wirecard noch an der Börse mit 23,6 Milliarden Euro bewertet. Diese waren am 25. Juni 2020 vernichtet, als der damalige DAX-Konzern Insolvenz angemeldet hat. 2022 begann der Strafprozess gegen einige Beteiligte, wie zum Beispiel Markus Braun, den damaligen Vorstandsvorsitzenden. Der damalige Finanzvorstand Jan Marsalek ist bis dato auf der Flucht. Ende Januar 2025 hat das Münchner Landgericht und die Staatsanwaltschaft den laufenden Strafprozess modifiziert und konzentriert sich auf die schwerwiegenden Vorwürfe. Markus Braun drohen bis zu 15 Jahre Haft.

2. Datenauswahl und Untersuchungszeitraum

Aufgrund der Brisanz und der Aktualität ist besonders zu erwähnen, dass die beschriebene Analyse ausschließlich auf öffentlich zugänglichen Informationen beruht. Mitte 2024 wurden alle Quartals- und Jahresberichte der Wirecard Gruppe von Quartal 1 aus dem Jahr 2006 bis einschließlich

Quartal 3 aus 2019 bei EQS heruntergeladen und archiviert.¹ Derzeit sind bei EQS keine Wirecard-Berichte mehr abrufbar.² Vereinzelt lassen sich noch einzelne Berichte über direkte Links abrufen.³ Auf der Unternehmens-Website fehlen folgende Berichte: Jahresberichte 2006 bis 2009, Quartal 1 bis Quartal 3 aus dem Jahr 2010, Quartal 1 bis Quartal 2 aus 2011 sowie Quartal 1 aus 2012.⁴

Der Untersuchungszeitraum beginnt ab 2006. Wirecard wurde ab dem 18. September 2006 in den TecDAX aufgenommen, und ab diesem Jahr waren die Finanzdaten vollständig zugänglich. Zu diesem Zeitpunkt war Markus Braun bereits seit 2002 als CEO der Wirecard AG tätig und prägte maßgeblich die strategische Ausrichtung des Unternehmens. Jan Marsalek, der seit 2000 für Wirecard arbeitete, übernahm 2010 die Position des Vertriebsvorstands und wurde 2017 schließlich zum Chief Operating Officer ernannt. Beide spielten eine zentrale Rolle in den mutmaßlich betrügerischen Aktivitäten, die später zur Insolvenz führten.

¹ Siehe <https://www.eqs.com> (Stand: 05.03.2025).

² Siehe <https://www.eqs-news.com/de/company/wirecard-ag/berichte/5b74e7bd-ea7c-11e8-902f-2c44fd856d8c> (Stand: 05.03.2025).

³ Siehe z. B. https://irpages2.eqs.com/download/companies/wirecard/Annual%20Reports/GB2008_WirecardAG_DE_.pdf (Stand: 05.03.2025).

⁴ Siehe <https://www.wirecard.com/de/finanzberichte/> (Stand: 05.03.2025).

Die in diesem Artikel gemachten Aussagen beziehen sich explizit auf den Gesamtuntersuchungszeitraum von 2006 bis 2019. Die Analyse kann selbstverständlich auch für zeitliche Teilabschnitte erfolgen.

3. Anwendung von Bilanzkennzahlen im Rahmen des Beneish-M-Score-Modells

Eine Möglichkeit des Forensic Accounting besteht darin, dass für die einzelnen Quartals- und Jahresberichte Bilanzkennzahlen berechnet werden, woraus sich ein Gesamtbild ergibt.

Dieser Ansatz wurde zunächst im Jahre 1998 von MBA-Studentinnen und -Studenten in der Klasse von Prof. Charles Lee an der Cornell Johnson Graduate School of Management gewählt. Die Studierenden kamen zu dem Schluss, dass die Aktien von Enron überbewertet waren und das Unternehmen möglicherweise den Gewinn manipuliert hatte. In diesem Zusammenhang wurde das sogenannte Beneish-M-Score-Modell entwickelt, was sich seitdem im Bereich Forensic Accounting etabliert hat.⁵ 2001 meldete Enron Insolvenz an.

Beim Beneish-M-Score-Modell werden die folgende acht Kennzahlen errechnet:

- Forderungen aus Lieferungen und Leistungen/ Umsatz,
- Veränderung des Bruttogewinns vom Umsatz,
- Veränderung der Qualität der Vermögenswerte,
- Umsatzwachstum,
- Veränderung der Abschreibungen,
- Veränderung der Vertriebs- und Verwaltungsaufwendungen,
- Veränderung des Verschuldungsgrads sowie
- Anteil nicht zahlungswirksamer Erträge.

Diese acht Kennzahlen werden unterschiedlich gewichtet und geben aggregiert an, ob mit einer gewissen Irrtumswahrscheinlichkeit von einem Bilanzbetrug ausgegangen werden kann. Im Fall Wirecard ergab sich für oben erwähnten Untersuchungszeitraum kein Verdacht auf Bilanzbetrug. Kritisch ist anzumerken, dass dem Beneish-M-Score-Modell zahlreiche Prämissen zugrunde liegen, die nicht auf Wirecard zutreffen:

- Wirecards GuV beruht auf dem Gesamtkostenverfahren und nicht dem Umsatzkostenverfahren.

- Somit sind die Vertriebs- und Verwaltungsaufwendungen bei Wirecard auch nicht separat ausgewiesen.
- Der konsolidierte Jahresabschluss der Wirecard-Gruppe beinhaltet die Wirecard Bank und unterscheidet sich somit signifikant von dem eines Industrieunternehmens.
- Bei Wirecard besteht deren langfristiges Vermögen überwiegend aus immateriellen Vermögenswerten wie Lizenzen, Software oder Rechten. Das Beneish-Modell berücksichtigt auf der anderen Seite hauptsächlich materielle Vermögensgegenstände.

Aus dem Blickwinkel der im Fall Wirecard notwendigen Änderung von Prämissen sind auch folgende Forschungsarbeiten zu sehen:

Spătăcean analysierte die Anwendung des Beneish-Modells zur Früherkennung von Bilanzbetrug und zeigte, dass bereits zwischen 2016 und 2018 auffällige Risikoindikatoren in den Finanzberichten von Wirecard erkennbar waren.⁶ Jo et al.⁷ untersuchten die Rolle der Unternehmensführung und der deutschen Finanzaufsicht im Fall Wirecard. Karami⁸ analysierte systemische Schwächen und das Versagen interner Kontrollmechanismen. Seine Untersuchung zeigt, dass Wirecard ein betrügerisches Netzwerk aus Tochtergesellschaften und fingierten Kunden aufbaute, das durch unzureichende Prüfungen und mangelnde Reaktionen auf Warnsignale begünstigt wurde. Die Untersuchung erfolgt im Rahmen des Beneish-Modells und der Benford Verteilung. Einschränkend ist anzumerken, dass beim Beneish-Modell zentrale Anpassungen (zum Beispiel Gesamtkostenverfahren anstatt Umsatzkostenverfahren) unterblieben und die Benford-Analyse lediglich auf den Umsatz- und Kassenpositionen beruht – bei einer Stichprobengröße von 56. In diesem Zusammenhang beleuchteten Löw und Heyd⁹ die Rolle der Abschlussprüfer von EY.

Neben klassischen Prüfmethode wurde auch der Einsatz moderner Technologien untersucht. Faccia et al.¹⁰ analysierten den Nutzen von Natural Language Processing zur Überprüfung der Finanzberichte von Wirecard und stellten fest, dass Sentiment-basiert Analysen potenzielle Hinweise auf Manipulationen liefern können. Allerdings seien sie nur in Kombination mit traditionellen Methoden eine verlässliche Betrugserkennung.

6 Vgl. Spătăcean, I. O. (2020).

7 Vgl. Jo, H. et al. (2021).

8 Vgl. Karami (2023).

9 Vgl. Löw, E./Heyd, R. (2024).

10 Vgl. Faccia, A. et al. (2024).

5 Siehe Beneish, M. D. (1999).

Schließlich beleuchteten Teichmann et al.¹¹ die weitreichenden Auswirkungen auf das Vertrauen in die deutschen Finanzmärkte und betonten die Notwendigkeit tiefgreifender Reformen, um ähnliche Fälle in Zukunft zu verhindern.

Leider liegt wohl aktuell noch kein Accounting-basiertes Modell vor, das oben genannte Einschränkungen aufhebt und somit für den Fall Wirecard mit einer sehr geringen Irrtumswahrscheinlichkeit auf einen Betrug hinweisen würde. Deshalb kommt hier dem auf Data Analytics basierenden Ansatz eine wichtige Rolle zu.

4. Data Analytics als wertvolles Tool im Fall Wirecard

Es wurde die folgende Hypothese formuliert: „In den GuV- und Bilanzkennzahlen von Wirecard finden sich signifikante Abweichungen im Vergleich zur erwarteten Verteilung, die auf mögliche Manipulationen hinweisen.“

Um diese Hypothese zu überprüfen, wurden die GuV- und Bilanzzahlen von Quartal 1 aus 2006 bis Quartal 3 aus 2019 herangezogen. Auf die Cashflow Statements wurde nicht zurückgegriffen, da diese Zahlen von den GuV- und Bilanzzahlen abhängig sind. Entsprechend wurden auch die abhängigen Finanzzahlen, wie zum Beispiel EBIT, Gewinn vor Steuern, Jahresüberschuss, Anlagevermögen, Umlaufvermögen oder Gesamtkapital, exkludiert.

4.1 Erwartete Verteilung

Im Rahmen der Hypothese ist die erwartende Verteilung wie folgt:

1. Die erste Ziffer sollte Benford-verteilt sein (die Wahrscheinlichkeit zum Beispiel für die Ziffer 1 als erste Ziffer beträgt 30,1 Prozent und für die Ziffer 9 4,6 Prozent).
2. Bei der zweiten Ziffer ergibt sich eine starke Glättung (die Wahrscheinlichkeit zum Beispiel für die Ziffer 1 beträgt 11,4 Prozent und für die Ziffer 9 8,5 Prozent).
3. Bei der dritten Ziffer ergibt sich fast eine Linearisierung (die Wahrscheinlichkeit zum Beispiel für die Ziffer 1 beträgt 10,2 Prozent und für die Ziffer 9 9,8 Prozent).
4. Die letzte Ziffer ist gleichverteilt (die Wahrscheinlichkeit zum Beispiel für die Ziffer 1 beträgt annähernd zehn Prozent und für die Ziffer 9 ebenfalls annähernd zehn Prozent).
5. Die ersten beiden Ziffern sollten sich analog zu Punkt 1 und 2 ergeben (die Wahrschein-

lichkeit, dass zum Beispiel die ersten beiden Ziffern 10, 11, 12, ... 18 oder 19 sind, ist 30,1 Prozent. Die Wahrscheinlichkeit für die 10 beträgt 4,1 Prozent und für die 19 2,2 Prozent).

Es deutet vieles darauf hin, dass Unternehmen einer bestimmten Größenordnung mit einer Prüfmethode nach dem Benfordschen Gesetz vertraut sein könnten.¹² Daher beschränkt sich diese Analyse nicht nur auf die Verteilung der ersten Ziffer, sondern die Analyse wurde auch auf die zweite, dritte und letzte Ziffer sowie die Kombination aus erster und zweiter Ziffer ausgeweitet.

4.2 Überprüfung der Abweichungen zur erwarteten Verteilung

Der hypothetischen Verteilung wurde die tatsächliche Verteilung der GuV- und Bilanzzahlen von Wirecard gegenübergestellt, und die Abweichungen wurden mit folgenden statistischen Tests überprüft:

- **Chi-Quadrat-Test:**
Dieser Test wird häufig eingesetzt, um die Übereinstimmung zwischen beobachteten und erwarteten Häufigkeiten in diskreten Daten zu bewerten. Insbesondere bei der Analyse von Finanzberichten, die dem Benfordschen Gesetz folgen sollten, hat sich der Chi-Quadrat-Test aufgrund seiner Flexibilität und Robustheit als verlässliches Werkzeug etabliert.¹³
- **Kolmogorow-Smirnow-Test:**
Er stellt in diesem Zusammenhang ein besonders geeignetes Werkzeug dar, da er die kumulative Verteilungsfunktion der beobachteten Daten mit der theoretischen Verteilung vergleicht. Der Kolmogorow-Smirnow-Test misst die maximale Differenz zwischen diesen Verteilungen und ermöglicht so eine präzise Identifikation von Abweichungen. Joenssen¹⁴ vergleicht verschiedene statistische Methoden zur Prüfung der Konformität mit dem Benfordschen Gesetz und kommt zu dem Ergebnis, dass der Kolmogorow-Smirnow-Test insbesondere bei mittelgroßen Abweichungen eine hohe Aussagekraft besitzt. Dies macht ihn zu einem zuverlässigen Instrument zur Identifikation potenzieller Manipulationen, insbesondere in komplexen Betrugsfällen wie dem Wirecard-Skandal.
- **Freedman-Watson U²-Test:**
Anders als der Kolmogorow-Smirnow-Test berücksichtigt Freedmans Methode nicht nur

¹² Vgl. Druică, E. et al., (2018).

¹³ Vgl. Joenssen, D. W. (2013).

¹⁴ Vgl. Joenssen, D. W. (2014).

¹¹ Vgl. Teichmann, F. et al. (2023).

Kolmogorow-Smirnow-Test Ergebnisse						
Signifikanzniv.	WIREFCARD	Adyen N.V.	Block Inc.	Paypal Inc.	Worldline S.A.	Peer Group
First Digit	5% *	✓	✓	*	✓	✓
Second Digit	5% *	✓	*	✓	✓	✓
Third Digit	5% *	✓	✓	✓	✓	✓
First Two Digits	5% *	✓	✓	*	✓	✓
Last Digit	5% *	✓	✓	✓	✓	✓

Freedman-Watson U-Quadrat-Test Ergebnisse						
Signifikanzniv.	WIREFCARD AG	Adyen N.V.	Block Inc.	Paypal Inc.	Worldline S.A.	Peer Group
First Digit	5% *	✓	✓	*	✓	✓
Second Digit	5% *	✓	*	✓	✓	✓
Third Digit	5% *	✓	✓	✓	✓	✓
First Two Digits	5% *	✓	✓	✓	✓	✓
Last Digit	5% *	✓	✓	✓	✓	✓

Chi-Quadrat-Test Ergebnisse						
Signifikanzniv.	WIREFCARD AG	Adyen N.V.	Block Inc.	Paypal Inc.	Worldline S.A.	Peer Group
First Digit	5% *	✓	✓	✓	✓	✓
Second Digit	5% *	✓	✓	✓	✓	✓
Third Digit	5% *	✓	✓	✓	✓	✓
First Two Digits	5% *	✓	*	✓	*	*
Last Digit	5% *	✓	✓	✓	✓	✓

Abb. 1: Die Untersuchungsergebnisse im Überblick

die maximalen kumulativen Abweichungen, sondern auch die systematischen Unterschiede zwischen den beobachteten und erwarteten Häufigkeiten sowie deren systematische Verteilungsmuster. Die Berechnung basiert auf der Differenz zwischen den empirischen und den theoretischen Häufigkeiten, wobei größere Abweichungen durch Quadrierung stärker gewichtet werden. Studien wie die von Qu, Steinberg und Burger¹⁵ belegen seine hohe Effektivität in der Aufdeckung von Anomalien in ökonomischen Daten.

Somit zeigt der Chi-Quadrat-Test einzelne Abweichungen auf, der Kolmogorow-Smirnow-Test die kumulierte Abweichung und der Freedman-Watson U^2 -Test die quadrierte Abweichung. Für den vorliegenden Fall scheint der Freedman-Watson U^2 -Test derjenige zu sein, der am schwersten manipuliert werden kann. In der vorliegenden Analyse wurden alle drei Tests durchgeführt. Es wurde von einem Signifikanzniveau von fünf Prozent ausgegangen. Zur Validierung der Ergebnisse wurden auch Szenarien für Signifikanzniveaus von ein Prozent und zehn Prozent berechnet, ohne dass sich hierbei signifikante Unterschiede bei den Ergebnissen zeigten.

4.3 Backtesting mit einer Peer Group

Um zu vermeiden, dass sich im Bereich von Zahlungsdienstleistern Anomalien ergeben, die im vorliegenden Fall fälschlicherweise als Betrug

interpretiert werden könnten, wurde ein Backtesting mit einer Peer Group vorgenommen. Als Vergleichsgruppe wurden die vier börsennotierten Zahlungsdienstleister Adyen, Block, Paypal und Worldline ausgewählt.¹⁶ Die für Wirecard durchgeführte Analyse wurde für alle vier Unternehmen einzeln und auch für die Peer Group insgesamt durchgeführt. Als Untersuchungszeitraum wurden aus Gründen der Praktikabilität die Jahre 2018 bis 2023 (Adyen und Worldline), 2017 bis 2023 (Paypal) sowie 2015 bis 2023 (Block) herangezogen. Abgesehen vom geänderten Untersuchungszeitraum wurde der Analysevorgang exakt analog zu Wirecard durchgeführt. Es ergaben sich folgende Ergebnisse (siehe Abbildung 1):

Abbildung 1 zeigt ein recht eindeutiges Bild. Im Fall Wirecard war die Abweichung der tatsächlichen Verteilung von der zu erwartenden Verteilung der Ziffern bei allen drei Testarten signifikant. Der einzige Fall, wo die Abweichung nicht signifikant war, war bei der Analyse der ersten beiden Ziffern im Rahmen des Freedman-Watson U^2 -Test. Somit kann in 14 von 15 untersuchten Fällen ein Bilanzbetrug vermutet werden. Die Zahlen von Adyen weisen keinerlei Abweichung zur erwarteten Verteilung auf. Bei Worldline und der gesamten Peer Group gibt es lediglich eine Abweichung bei den ersten beiden Ziffern im Rahmen des Chi-Quadrat-Tests. Block und Paypal weisen jeweils Abweichungen in drei Fällen auf.

¹⁵ Vgl. Qu, H./Steinberg, R./Burger, R. (2019).

¹⁶ Zur Auswahl einer Peer Group vgl. Ernst, D./Häcker, J. (2011).

4.4 Einschränkungen der Ergebnisse

Die in Abbildung 1 dargestellte Aussage ist nur so gut, wie die der Analyse zugrunde liegenden Prämissen. In einem weiteren Schritt sind gegebenenfalls die einschränkenden Prämissen noch zu erweitern.

1. Es wurden insgesamt 1.950 Zahlen in die Analyse von Wirecard einbezogen. Bei der gesamten Peer Group waren es 1.064. Selbstverständlich könnte die Anzahl noch vergrößert werden, wenn noch umfangreicheres Datenmaterial zur Verfügung stünde. Zusätzlich könnten die Betrachtungszeiträume der Peer Group, wenn möglich, analog zu den Zeiträumen von Wirecard gewählt werden.
2. Alle Zahlen wurden gleichgewichtet. Ein potenzieller Betrug wird jedoch eher beispielsweise in der Umsatzzeile oder bei den Forderungen aus Lieferungen und Leistungen stattfinden, als bei den sonstigen betrieblichen Erträgen. Hier zeigt sich wiederum die Notwendigkeit, das auf Industrieunternehmen ausgelegte Beneish-M-Score-Modell auf Unternehmen zu erweitern, die hohe immaterielle Vermögensgegenstände besitzen und Bankgeschäft betreiben.
3. Wirecard hat vor Quartal 3 des Jahres 2012 die Zahlen Cent-genau aufgeführt. Aufgrund des starken Wachstums wurde danach in 1.000 Euro und ab Quartal 2 des Jahres 2018 in 1.000.000 Euro berichtet. Bei der Analyse der letzten Ziffer wurden ab Quartal 3 des Jahres 2012 die letzten Nullen ignoriert und als letzte Ziffer der einstellige Eurobetrag herangezogen.
4. Aus der Gesamtheit der in die Analyse einbezogenen Zahlen lässt sich sowohl eine GuV als auch eine Bilanz und ein Cashflow Statement ableiten. Abgesehen von Summenpositionen, wie zum Beispiel dem EBIT, wurden keine weiteren Zahlen exkludiert. Alle Inputzahlen wurden einer Kontrolle unterzogen. So wurde beispielsweise der EBIT nicht als Zahl übernommen, sondern alle Aufwandspositionen wurden vom Umsatz subtrahiert. Das Ergebnis wurde dann mit dem ursprünglich importierten EBIT verglichen. Da das EBIT jedoch eine abhängige Variable ist, wurde die Position lediglich als Kontrollposition herangezogen und nicht weiter in die Untersuchung mit einbezogen.
5. Innerhalb der Wirecard-Gruppe ist die Wirecard Bank AG konsolidiert. Damit hat der

Bilanzersteller innerhalb der gesetzlichen Rahmenbedingungen mehr Gestaltungsspielräume (zum Beispiel 340f HGB). Die vorgenommenen Veränderungen müssen jedoch im Anhang erläutert werden. Sollten Berichtigungen zu Vorjahren vorgenommen werden, so entsprechen diese jedoch ebenfalls oben aufgezeigter Verteilung. Sollten Abweichungen in betrügerischer Absicht vorgenommen worden sein, so weicht die tatsächliche Verteilung der entsprechenden Ziffer von der erwarteten Verteilung ab.

6. Als Vergleichsgruppe wurden vier Unternehmen herangezogen. Die Peer Group könnte noch vergrößert werden, und die Analyse könnte auch nicht nur Zahlungsdienstleister miteinbeziehen.

Literaturverzeichnis

- Druică, E./Oancea, B./Vălsan, C. (2018): Benford's law and the limits of digit analysis, *International Journal of Accounting Information Systems*, 31, S. 75–82.
- Ernst, D./Häcker, J. (2011): *Applied International Corporate Finance*, Vahlen Verlag, München.
- Faccia, A./McDonald, J./George, B. (2024): NLP sentiment analysis and accounting transparency: A new era of financial record keeping, *Computers*, 13(5).
- Jo, H./Hsu, A./Llanos-Popolizio, R./Vergara-Vega, J. (2021): Corporate governance and financial fraud of Wirecard, *European Journal of Business and Management Research*, 6(2), S. 96–115.
- Joenssen, D.W. (2013): Two-digit testing for Benford's Law, *Proceedings 59th ISI World Statistics Congress*, Hong Kong, 25–30 August 2013, S. 3881–3886.
- Joenssen, D.W. (2014): Testing for Benford's Law, A Monte Carlo Comparison, *SSRN Electronic Journal*, <https://doi.org/10.2139/ssrn.2545243> (Stand: 05.03.2025).
- Karami, B. (2022): *Skandalfall Wirecard, Eine wissenschaftlich-fundierte interdisziplinäre Analyse*, Springer Gabler Wiesbaden.
- Löw, E./Heyd, R. (2024): *The Audit Failures of the Wirecard Scandal*, Palgrave Macmillan Cham.
- Qu, H./Steinberg, R./Burger, R. (2019): Abiding by the Law? Using Benford's Law to Examine the Accuracy of Nonprofit Financial Reports, *Nonprofit and Voluntary Sector Quarterly*, 48(4), S. 1–16.
- Spătacean, I. O. (2020): Testing early signs of fraudulent financial reporting – The Wirecard case, *Acta Marisensis, Seria Oeconomica*, 20(1), S. 19–35.
- Teichmann, F./Boticiu, S. R./Sergi, B. S. (2023): Wirecard scandal, A commentary on the biggest accounting fraud in Germany's post-war history, *Journal of Financial Crime*, 30(1), S. 1–8.