



Textauswertung in der Finanzanalyse: Eine Ereignisstudie zum Sentiment in Earnings Calls

Leonie Terhardt

Heinrich Heine University Düsseldorf

Abstract

This thesis examines the influence of sentiment in earnings calls on financial market reactions. The objective is to analyse the extent to which the tone used in presentations and Q&A segments affect abnormal returns, and how differences in investors' information environments shape these effects. The study is based on transcripts from companies in the S&P Composite 1500, whose tone is recorded using finance-specific lexicons and evaluated within an event-study framework. The results suggest that sentiment in the Q&A section in particular triggers significant immediate market reactions, with analysts' questions having the greatest influence. Conversely, the sentiment of the presentation assumes significance during the drift period and frequently results in subsequent corrections. In addition, investors show increased sensitivity to verbal signals when it comes to companies with a higher degree of uncertainty. The thesis contributes to the literature by demonstrating that context-specific sentiment measurement provides valuable insights for the evaluation of qualitative company information. Furthermore, it highlights the limitations of generic text lexicons.

Zusammenfassung

Die vorliegende Arbeit untersucht den Einfluss des Sentiments in Earnings Calls auf die Reaktionen der Finanzmärkte. Ziel ist es zu analysieren, inwiefern der sprachliche Ausdruck in der Präsentation und im Rahmen des Q&A-Segments abnormale Renditen beeinflusst und welche Bedeutung die Informationslage der Investoren dabei hat. Die Untersuchung basiert auf Transkripten von Unternehmen des S&P Composite 1500, deren Tonalität mithilfe finanzspezifischer Lexika erfasst und im Rahmen einer Ereignisstudie ausgewertet wird. Die Ergebnisse der Untersuchung legen nahe, dass insbesondere das Sentiment im Q&A-Teil signifikante unmittelbare Marktreaktionen auslöst, wobei die Fragen der Analysten den größten Einfluss besitzen. Demgegenüber gewinnt das Sentiment der Präsentation in der Driftperiode an Bedeutung und führt häufig zu späteren Korrekturen. Darüber hinaus zeigen Investoren in Bezug auf Unternehmen mit einer höheren Unsicherheit eine erhöhte Sensibilität auf verbale Signale. Die Arbeit leistet einen Beitrag zur Forschung, indem sie zeigt, dass eine kontextspezifische Sentimentmessung zentrale Einblicke für die Bewertung qualitativer Unternehmensinformationen liefert und dabei die Grenzen generischer Textlexika aufzeigt.

Keywords: sentiment; earnings calls; textanalyse; ereignisstudie; finanzmarktreaktionen

1 Einleitung

Während traditionell finanzielle Kennzahlen wie die Gewinne pro Aktie im Fokus standen, wird in der Finanzanalyse zunehmend erkannt, dass diese allein nicht ausreichen, um ein umfassendes Bild der Unternehmensleistung zu erhal-

ten.¹ Für die Einschätzung von Chancen, Risiken sowie der strategischen Unternehmensausrichtung sind vor allem qualitative Informationen aus Unternehmensveröffentlichungen von Bedeutung. Vor diesem Hintergrund wächst das Interesse an der systematischen Auswertung sprachlicher Inhal-

¹ Vgl. Davis und Tama-Sweet (2012, p. 806); Feldman et al. (2009, p. 916).

te der Unternehmenskommunikation. Ein besonders vielversprechender Ansatz in diesem Bereich ist die Sentimentanalyse, die es ermöglicht, Stimmungen in Texten zu quantifizieren.² Insbesondere Earnings Calls – Telefonkonferenzen, in denen Unternehmen ihre finanziellen Ergebnisse präsentieren und sich den Fragen von Analysten stellen – erweisen sich hierbei als wertvolle Informationsquelle.³ Die dynamische Interaktion, insbesondere im Frage- und Antwortabschnitt, liefert oft wertvolle qualitative Informationen, die über die rein formalen Berichterstattungen hinausgehen und subtile Hinweise auf die Stimmung der Unternehmensvertreter und Analysten sowie auf die Zukunftsaussichten des Unternehmens geben.

Die vorliegende Arbeit zielt darauf ab, aufzuzeigen, wie die Textauswertung als ergänzendes Instrument zur klassischen Finanzanalyse eingesetzt werden kann. Konkret wird untersucht, welche Rolle die Sentimentanalyse im Kontext von Earnings Calls spielt, um ein tieferes Verständnis für den Zusammenhang zwischen den in Unternehmensmitteilungen ausgedrückten Stimmungen und dem Marktverhalten zu erlangen. Angesichts widersprüchlicher Forschungsergebnisse soll insbesondere geklärt werden, wie das Sentiment in verschiedenen Abschnitten von Earnings Calls die Marktreaktionen beeinflusst und welche Rolle dabei die Informationslage der Anleger sowie methodische Unterschiede in der lexikonbasierten Sentimentanalyse spielen.

Hierzu ist die Arbeit folgendermaßen aufgebaut: Im zweiten Kapitel werden die theoretischen Grundlagen und der aktuelle Forschungsstand der textbasierten Sentimentanalyse sowie deren Anwendung auf Earnings Calls dargelegt. Dabei wird die Bedeutung qualitativer Inhalte in Earnings Calls als wesentliche Informationsquelle erörtert und aufgezeigt, wie das in Texten enthaltene Sentiment ausgewertet werden kann. Das dritte Kapitel beschreibt die methodische Vorgehensweise. Anhand einer Stichprobe von Earnings Calls aus dem S&P Composite 1500 werden mit einem lexikonbasierten Ansatz Sentiment-Indikatoren für verschiedene Abschnitte und Redner ermittelt. Um die Marktreaktionen auf das Sentiment der Earnings Calls zu analysieren, werden darüber hinaus mittels einer Ereignisstudien-Methodik abnormale Renditen berechnet. Im vierten Kapitel werden die Ergebnisse der Auswirkungen des Sentiments in verschiedenen Segmenten von Earnings Calls auf die Marktreaktionen dargestellt und diskutiert. Ein besonderer Fokus liegt hierbei auf den Unterschieden, die sich je nach verwendetem Lexikon zur Ermittlung ergeben, um die Sensitivität gegenüber unterschiedlichen textanalytischen Methoden zu beleuchten. Darüber hinaus wird der Einfluss der Informationslage auf die Reaktionen der Investoren und die Robustheit der Ergebnisse überprüft. Die Arbeit schließt im fünften Kapitel mit einer Zusammenfassung der wichtigsten Erkenntnisse sowie Empfehlungen für zukünftige Forschungsarbeiten ab.

Die Ergebnisse der Arbeit bestätigen, dass emotionale

und sprachliche Nuancen aus Earnings Calls eine zentrale Rolle spielen und mithilfe der Sentimentanalyse systematisch erfasst werden können. Hierbei hat der Q&A-Abschnitt, insbesondere die Fragen der Analysten, einen besonders starken Einfluss auf die Marktreaktionen. Nach einer anfänglich emotional geprägten Reaktion beginnen Investoren, die langfristigen Implikationen der erhaltenen Informationen genauer zu analysieren, was in der Driftperiode häufig zu einer negativen Korrektur der anfänglich positiven Marktreaktionen führt. Die Reaktionen der Anleger hängen dabei stark von ihrer Informationslage ab, wobei Unternehmen mit höherer Unsicherheit stärkere Marktreaktionen auf das Sentiment zeigen. Erfahrene Investoren sind besser in der Lage, wertrelevante Informationen aus dem Ton von Earnings Calls zu extrahieren, während unerfahrene Anleger ihre initiale Reaktion häufiger revidieren. Der Einfluss der einzelnen Abschnitte und Redner eines Earnings Calls variiert ebenfalls je nach Informationslage. Insbesondere die Antworten des Managements im Q&A-Abschnitt gewinnen bei zunehmenden Informationsasymmetrien an Bedeutung. Auch erfahrene Anleger messen ihnen eine größere Relevanz bei. Die Arbeit unterstreicht zudem die methodischen Herausforderungen einer lexikonbasierten Sentimentanalyse bei der Erfassung von Kontext und Mehrdeutigkeit. Hierbei wird insbesondere deutlich, dass für eine präzise Auswertung der Texte jedes Segment eines Earnings Calls ein spezifisch angepasstes Lexikon erfordert, und nicht auf ein universelles Lexikon zurückgegriffen werden sollte.

2 Theoretische Grundlagen der Sentimentanalyse in Earnings Calls

2.1 Earnings Calls als Medium der qualitativen Offenlegung

Im zweiten Quartal 2019 veröffentlichte Netflix Inc. starke Umsatz- und Gewinnzahlen, die auf den ersten Blick eine positive Entwicklung des Unternehmens signalisierten.⁴ Während des Earnings Calls gab das Unternehmen jedoch bekannt, dass die Anzahl neuer Abonnenten hinter den Erwartungen zurückblieb und äußerte Bedenken hinsichtlich des zukünftigen Wachstums durch zunehmenden Wettbewerbsdruck. Diese qualitativen Informationen führten trotz positiver Finanzkennzahlen zu einem Kursverlust von über 10 %.

Das angeführte Beispiel verdeutlicht die zentrale Bedeutung qualitativer Informationen in der Finanzkommunikation und deren potenziellen Einfluss auf die Marktdynamik, unabhängig davon, ob die quantitativen Ergebnisse positiv oder negativ ausfallen. Die finanzielle Berichterstattung eines Unternehmens zielt darauf ab, Informationsasymmetrien und Interessenkonflikte zwischen Managern und externen Stakeholdern zu reduzieren.⁵ Sie umfasst quantitative Kennzahlen und qualitative Einschätzungen, sowie sowohl verpflichtende, als auch freiwillige

² Vgl. Todd et al. (2024, p. 2).

³ Vgl. De Amicis et al. (2021); Blau et al. (2015, p. 204); Brockman et al. (2015, pp. 24f).

⁴ Vgl. Rodriguez (2019).

⁵ Vgl. Davis und Tama-Sweet (2012, p. 804).

Offenlegungen. In der Vergangenheit hat sich Finanzanalyse primär auf numerische Aspekte konzentriert.⁶ Dies ist darauf zurückzuführen, dass quantitative Daten im Vergleich zu qualitativen leichter verfügbar, objektiver und eindeutiger sind. Zahlen allein können jedoch das gesamte Spektrum der Unternehmensleistung nicht vollständig abbilden.⁷ Qualitative Informationen, insbesondere die Stimmungen und Einschätzungen des Managements, spielen eine entscheidende Rolle bei der Interpretation von Chancen, Risiken und der allgemeinen Unternehmensstrategie. Die Integration von Aspekten der qualitativen Finanzberichterstattung in die Finanzanalyse eröffnet daher die Möglichkeit, ein ganzheitlicheres Bild der Finanzmarktdynamik zu zeichnen und hilft Investoren fundiertere Entscheidungen zu treffen.⁸

Die zunehmende Bedeutung qualitativer Informationen in der Finanzberichterstattung wird insbesondere anhand der Entwicklung sogenannter Earnings Calls deutlich. Hierbei handelt es sich um eine Präsentation und Diskussion der Finanzergebnisse eines Unternehmens durch dessen Management in Form einer Telefonkonferenz.⁹ Diese Konferenzen finden auf freiwilliger Basis kurz nach der Veröffentlichung der jeweiligen Quartals- oder Jahresergebnisse statt. Das Inkrafttreten der Regulation Fair Disclosure (Reg FD) im Jahr 2000 hat zu einer deutlich stärkeren Nutzung dieser Plattform durch Unternehmen für die Kommunikation mit Investoren geführt und den Informationsgehalt der Telefonkonferenzen erhöht.¹⁰ Die Reg FD verpflichtet börsennotierte Unternehmen dazu, wichtige, bislang nicht öffentlich bekannte, Informationen zu veröffentlichen. Vor dieser Regulierung waren Earnings Calls nur für ausgewählte große Investoren und Analysten über private Telefonnummern zugänglich.¹¹ Seit Einführung der Reg FD muss die Teilnahme auch der breiten Öffentlichkeit ermöglicht werden. Gegenwärtig erfolgt die Übertragung von Earnings Calls vielfach live über Investor-Relations-Websites oder Streaming-Dienste wie YouTube. Die simultane Übermittlung von Informationen gewährleistet eine gerechte Informationsverteilung und löst damit das Problem selektiver Offenlegungen.¹²

Im Rahmen der initialen Forschungsarbeiten zum Stellenwert von Earnings Calls erfolgte zunächst eine Untersuchung der grundlegenden Nutzung dieser Kommunikationsform, da seinerzeit nur eine unzureichende Datenbasis zu den tatsächlich kommunizierten Inhalten verfügbar war.¹³ Frankel et al. (1999) identifizierten Telefonkonferenzen als einflussreiches Informationsereignis, das durch signifikant hohe Renditevolatilität und Handelsvolumen gekennzeichnet ist.¹⁴ In einer weiteren Untersuchung zeigten Bowen et al. (2002), dass

die Prognosegenauigkeit von Analysten nach Gewinnankündigungen mit Earnings Calls im Vergleich zu solchen ohne Earnings Call zunimmt.¹⁵ Kimbrough (2005) konnte darüber hinaus nachweisen, dass der Markt auf aktuelle Gewinne weniger sensibel reagiert, wenn eine Gewinnankündigung mit einer Telefonkonferenz einhergeht, als wenn dies nicht der Fall ist.¹⁶

Ein Earnings Call beginnt typischerweise mit einer sogenannten „Safe Harbour“-Erklärung des Managements, in der darauf hingewiesen wird, dass die präsentierten Daten zukunftsgerichtete Aussagen enthalten, die von den tatsächlichen Ergebnissen abweichen können.¹⁷ Earnings Calls umfassen eine Präsentation sowie einen Frage-und-Antwort-Abschnitt (Q&A).¹⁸ Im Rahmen der Präsentation, die in der Regel eine Dauer von 15 bis 20 Minuten hat, stellen die Unternehmensvertreter die Ergebnisse des abgelaufenen Berichtszeitraums dar. In der Regel übernimmt der Chief Executive Officer (CEO) die Berichterstattung über die allgemeine Geschäfts- und Marktperformance sowie die strategische Vision, während der Chief Financial Officer (CFO) für die Erläuterung der finanziellen Details und Hintergründe zuständig ist. Weitere Führungskräfte können spezifische Einblicke in ihre Verantwortungsbereiche geben. Das anschließende Q&A erstreckt sich über einen Zeitraum von 30 bis 45 Minuten und bietet Analysten und anderen Teilnehmern die Möglichkeit, die präsentierten Informationen zu hinterfragen oder zusätzliche Details zu verlangen, die nicht in der Präsentation oder in anderen Offenlegungskanälen wie Geschäftsberichten oder Pressemitteilungen deutlich geworden sind. Die Unternehmensleitung kann durch gezielte Anfragen seitens der Teilnehmer zur Preisgabe privater Informationen veranlasst werden, einschließlich Angaben, die von den Managern lieber zurückbehalten werden würden. Manager sind zwar nicht verpflichtet, auf direkte Anfragen von Analysten zu antworten, jedoch besteht die Möglichkeit, dass der Markt aus einer Weigerung, Informationen zu veröffentlichen, negative Schlussfolgerungen zieht.¹⁹

Im Gegensatz zu standardisierten und vorbereiteten Dokumenten wie Finanzberichten oder Pressemitteilungen ist das Informationsumfeld einer Telefonkonferenz demnach dynamischer und unstrukturierter.²⁰ Insbesondere der spontane Charakter des Q&A-Abschnitts stellt eine Herausforderung für die Manager dar, da sie gleichzeitig den Inhalt und die Sprache ihrer Aussagen kontrollieren müssen.²¹ Über die inhaltlich neuen Offenlegungen hinaus bieten die qualitativen Aspekte dieser Gespräche potenziell informative verbale Hinweise, die auch emotionale Ausdrücke wie „zufrieden“ oder „ernüchtert“ umfassen können.²² Die Wortwahl

⁶ Vgl. Feldman et al. (2009, p. 916).

⁷ Vgl. Davis und Tama-Sweet (2012, p. 806).

⁸ Vgl. Feldman et al. (2009, p. 916).

⁹ Vgl. CFI Education Inc. (2024).

¹⁰ Vgl. Brockman et al. (2015, p. 24); Irani (2004, p. 26).

¹¹ Vgl. Elliott et al. (2023, p. 11).

¹² Vgl. Hollander et al. (2010, p. 536); Frankel et al. (1999, p. 135).

¹³ Vgl. Matsumoto et al. (2011, p. 1386).

¹⁴ Vgl. Frankel et al. (1999, pp. 134f).

¹⁵ Vgl. Bowen et al. (2002, pp. 285f).

¹⁶ Vgl. Matsumoto et al. (2011, p. 1386).

¹⁷ Vgl. CFI Education Inc. (2024).

¹⁸ Vgl. Hollander et al. (2010, p. 537).

¹⁹ Vgl. Hollander et al. (2010, p. 559).

²⁰ Vgl. Blau et al. (2015, p. 206).

²¹ Vgl. De Amicis et al. (2021).

²² Vgl. Elliott et al. (2023, p. 8); Matsumoto et al. (2011, p. 1384).

der Teilnehmer kann stark nuanciert sein, da Manager versuchen, das Unternehmen positiv darzustellen, ohne negative Konsequenzen zu riskieren.²³ Im Vergleich zu Zahlen bieten Worte Managern bei qualitativen Offenlegungen dabei einen größeren Spielraum. Obgleich sie sich der Problematik bewusst sind, offensichtlich falsche Aussagen zu treffen, ist die Formulierung von Faktenbehauptungen in positiveren oder weniger positiven Ausdrücken mit einem geringeren Risiko für regulatorische Interventionen verbunden.²⁴ Earnings Calls werden von Managern als weniger formell und haftungsintensiv wahrgenommen als Pressemitteilungen.²⁵ Infolgedessen sind sie während dieser Telefonkonferenzen eher geneigt, bestimmte Informationen preiszugeben, die sie in einer Pressemitteilung zurückhalten würden. Insbesondere dann, wenn spezifische Details in den formelleren Berichterstattungen nicht offengelegt werden können, kann der Tonfall in diesen Konferenzen eine Signalwirkung entfalten.²⁶ Earnings Calls zeichnen sich demnach durch eine subtile Informationsvermittlung aus, die durch traditionelle Methoden der Finanzanalyse nur schwer zu erfassen und auszuwerten ist.²⁷

2.2 Sentimentanalyse im Kontext von Markteffizienz und Behavioural Finance

Der Informationsgehalt von Earnings Calls wirft Fragen bezüglich der Markteffizienz auf, wie sie in der Effizienzmarkthypothese (EMH) beschrieben wird. Gemäß der EMH reflektieren die Preise von Vermögenswerten ihren fundamentalen Wert und enthalten sämtliche relevante Informationen, die zu einem gegebenen Zeitpunkt von rationalen Marktteilnehmern effizient und fehlerfrei genutzt werden.²⁸ Die EMH wird in drei Formen unterteilt: die starke, die semi-starke und die schwache Form. Die starke Form besagt, dass alle Informationen, sowohl öffentliche als auch private, in den Preisen enthalten sind. Dies impliziert, dass selbst Insider keinen Vorteil haben und keine überdurchschnittlichen Renditen erzielen können, da alle Informationen bereits vollständig im Preis integriert sind. Demzufolge sollte die Veröffentlichung von Earnings Calls keinen signifikanten Einfluss auf die Aktienkurse haben. Treten dennoch um Earnings Calls herum außergewöhnliche Renditen auf, lässt dies den Schluss zu, dass die Märkte nicht vollständig effizient sind und Informationen nicht unverzüglich oder vollständig eingepreist wurden. Die semi-starke Form der EMH modifiziert die ursprüngliche Hypothese, indem sie annimmt, dass sämtliche öffentlich verfügbaren Informationen – darunter vergangene Preisbewegungen, Unternehmensberichte und Nachrichten – im aktuellen Preisniveau abgebildet sind.²⁹ Es wird also angenommen,

dass Anleger auf neu veröffentlichte, unternehmensspezifische Informationen reagieren.³⁰ Dies eröffnet die Möglichkeit, dass sie ihre Erwartungen und Anlageentscheidungen basierend auf den Signalen in Earnings Calls anpassen, sofern diese über bereits bekannte oder antizipierte Informationen hinausgehen.³¹ In ihrer schwachen Form besagt die EMH demgegenüber lediglich, dass alle historischen Informationen in den aktuellen Preisen reflektiert sind.³²

Im Gegensatz zur EMH stellt die Behavioral Finance die Annahme der vollkommen rationalen Marktteilnehmer infrage und integriert psychologische Erkenntnisse in die Untersuchung von Finanzmärkten.³³ Sie nimmt an, dass die Preisbildungsmechanismen nicht nur von historischen, öffentlichen und privaten Informationen, sondern auch stark von individuellen Ansichten und Emotionen der Anleger beeinflusst werden. Soziale, verhaltensbedingte und psychologische Faktoren können die Erwartungen und Entscheidungen von Menschen maßgeblich beeinflussen, sodass allgemein bekannte Informationen nicht rational verarbeitet werden. Emotionen und kognitive Verzerrungen können demnach zu einer temporären Fehlbewertung des Werts von Vermögenswerten führen,³⁴ wobei Unsicherheit als verstärkender Faktor wirkt.³⁵ Die von der Behavioral Finance beschriebenen Verzerrungen im Anlegerverhalten stehen im Widerspruch zur EMH, gemäß derer Abweichungen zwischen Marktpreisen und ihren fundamentalen Werten durch rationale Marktteilnehmer, sogenannte Arbitrageure, zeitnah korrigiert und irrationale Händler vom Markt verdrängt werden.³⁶ Es konnte jedoch festgestellt werden, dass rationale Händler in der Regel risikoscheu sind, über einen kurzen Handelshorizont verfügen und geringere Renditen erzielen.³⁷ In der Konsequenz des begrenzten Handlungsspielraums rational handelnder Investoren kann es somit durch das Verhalten irrationeller Investoren zu Marktineffizienzen und fehlerhaften Preisbildungen am Kapitalmarkt kommen.³⁸

Vor diesem Hintergrund nimmt die Untersuchung der Auswirkung von Sentiment auf finanzielle Entscheidungen in der wissenschaftlichen Forschung zunehmend an Bedeutung zu.³⁹ Der Begriff „Sentiment“ wird im allgemeinen Sprachgebrauch verwendet, um Gefühle und Überzeugungen gegenüber einem bestimmten Sachverhalt oder einer Situation zu beschreiben⁴⁰ und wird in der deutschsprachigen Literatur häufig mit „Stimmung“ übersetzt.⁴¹ Die Sentimentanalyse zielt darauf ab, Meinungen, Einstellungen

²³ Vgl. Blau et al. (2015, p. 204).

²⁴ Vgl. Davis und Tama-Sweet (2012, p. 644).

²⁵ Vgl. Matsumoto et al. (2011, p. 1387)

²⁶ Vgl. Pouryousof et al. (2023).

²⁷ Vgl. Blau et al. (2015, p. 204).

²⁸ Vgl. Malkiel und Fama (1970, p. 383).

²⁹ Vgl. Malkiel und Fama (1970, p. 383).

³⁰ Vgl. Blau et al. (2015, p. 203).

³¹ Vgl. Rahman (2023).

³² Vgl. Malkiel und Fama (1970, p. 383).

³³ Vgl. Baddeley (2018, pp. 9, 14).

³⁴ Vgl. H. K. Baker und Nofsinger (2010, p. 340).

³⁵ Vgl. Baddeley (2018, p. 5).

³⁶ Vgl. H. K. Baker und Nofsinger (2010, p. 595).

³⁷ Vgl. Gregory (2021).

³⁸ Vgl. M. Baker und Wurgler (2007, p. 131).

³⁹ Vgl. M. Baker und Wurgler (2007, p. 131).

⁴⁰ Vgl. Liu (2020, p. 2).

⁴¹ Vgl. boerse.de (n. d.).

und Emotionen zu identifizieren und zu quantifizieren.⁴² In finanziellen Analysen fungiert sie als Bindeglied zwischen der Effizienzmarkthypothese und der Behavioral Finance, indem sie die Stimmungslage der Marktteilnehmer untersucht. Auf diese Weise lässt sich aufzeigen, wie emotionale und kognitive Verzerrungen die Märkte beeinflussen und zu Abweichungen von der rationalen Marktlogik der Effizienzmarkthypothese führen. Die gegenwärtige wissenschaftliche Literatur beschäftigt sich daher intensiv mit der Messung und Analyse des Einflusses von Sentiment auf das Verhalten von Marktteilnehmern sowie dessen Bedeutung für die Prognosefähigkeit realer Entwicklungen auf Finanzmärkten.⁴³

Der Begriff "Manager-Sentiment" bezeichnet die emotionalen und psychologischen Haltungen, die Manager in ihren Entscheidungen und Äußerungen zum Ausdruck bringen.⁴⁴ Diese können auch einen Verhaltensbias umfassen, der sich in irrationalen Überzeugungen über die zukünftige Aussicht ihrer Unternehmen manifestiert.⁴⁵ Im Rahmen ihrer Tätigkeit sammeln Manager private Informationen über die zukünftig erwarteten Cashflows und Risiken ihres Unternehmens.⁴⁶ Diese bleiben in den quantitativen Unternehmensberichten allerdings weitgehend unberücksichtigt, da solche Berichte primär vergangenheitsbezogene und standardisierte Daten enthalten, die nur geringe Einblicke in die internen Einschätzungen und strategischen Überlegungen des Managements bieten. Dahingegen können qualitative Offenlegungen den Anlegern zumindest teilweise dabei behilflich sein, das Manager-Sentiment und die damit verbundenen Insiderinformationen der Unternehmensleitung zu identifizieren und diese Informationen in ihre Anlageentscheidungen einfließen zu lassen. Earnings Calls stellen hierbei eine zentrale Form der Offenlegung dar, in denen subjektive Wahrnehmungen und emotionale Reaktionen zum Ausdruck kommen, die das Sentiment der Manager widerspiegeln.⁴⁷ Die präsentierten verbalen und nonverbalen Hinweise der Unternehmensführung können als Indikatoren für zukünftige Geschäftserwartungen betrachtet werden, die sich nicht unmittelbar aus den reinen Zahlen ableiten lassen. Daher wird folgende Hypothese aufgestellt:

H1a: *Ein positives Sentiment in Earnings Calls führt zu signifikant positiven Marktreaktionen, während ein negatives Sentiment signifikant negative Marktreaktionen hervorruft*

Sowohl das Sentiment von Managern als auch das von Investoren stellt einen wesentlichen Faktor in der Dynamik der Finanzmärkte dar.⁴⁸ Dennoch lassen sich Unterschiede in den Hauptinteressen und Einflussfaktoren feststellen. Investoren fokussieren sich primär auf die erwarteten Renditen, Wachstumsaussichten und Risiken in ihren Portfolios. Ihr

Sentiment umfasst die kollektiven emotionalen und psychologischen Haltungen gegenüber einem Unternehmen oder dem Markt im Allgemeinen, die nicht durch wirtschaftliche Fundamentaldaten gestützt sein müssen. Im Gegensatz dazu ist die Unternehmensführung primär an den zukünftigen Aussichten der von ihnen betreuten Unternehmen interessiert. Gregory (2021) konnte nachweisen, dass die Prognosefähigkeit des Manager-Sentiments in erster Linie auf privaten Informationen basiert, während die Investorenstimmung überwiegend durch öffentliche Informationen geprägt ist.⁴⁹ Dies zeigt, dass beide Gruppen über unterschiedliche Informationsgrundlagen verfügen. Darüber hinaus stellten Jiang et al. (2019) fest, dass das Sentiment von Managern und Investoren nicht in gegenseitiger Abhängigkeit zueinandersteht, sondern unterschiedliche Aspekte des Marktsentiments erfasst und sich ergänzt.⁵⁰ Sie betonen, dass das Manager-Sentiment stark durch interne Erwartungen getrieben wird, die häufig übermäßig optimistisch sind.

Empirische Evidenz sowie theoretische Überlegungen legen nahe, dass insbesondere Aktien, deren Bewertung schwierig ist, deren Handel kostenintensiv ist oder die einen spekulativen Charakter besitzen, eine höhere Anfälligkeit für Sentiment aufweisen.⁵¹ Die Interpretation von Informationen gestaltet sich für Investoren mit unzureichendem Informationsstand als Herausforderung. Dies kann dazu führen, dass vermehrt auf alternative Informationsquellen zurückgegriffen wird, um die bestehende Informationsasymmetrie zu reduzieren. Auch bei der Untersuchung von Earnings Calls wird ersichtlich, dass Informationsasymmetrien zu einer verstärkten Beeinflussung durch Sentiment führen. Bochkay et al. (2020) zeigten beispielsweise, dass bei Unternehmen mit einem weniger transparenten Informationsumfeld sowohl das Handelsvolumen als auch die Kursreaktionen während einer Telefonkonferenz ansteigen, wenn sich die Manager einer extremeren Sprache bedienen.⁵² Auf Basis dieser Erkenntnisse lässt sich folgende Hypothese formulieren:

H1b: *Je geringer die Informationslage der Investoren eines Unternehmens, desto stärker reagieren die Investoren auf das Sentiment in Earnings Calls*

Diese Differenzierung in der Wahrnehmung des Sentiments führt zu der Frage, wie Manager durch gezielte Offenlegungen und Sprachstrategien versuchen, Informationsasymmetrien zu verringern und die Erwartungen der Investoren zu steuern. Gemäß der Signaling-Theorie nutzen Manager freiwillige Offenlegungen und den Sprachstil ihrer Aussagen, um die Erwartungen der Investoren bezüglich künftiger Unternehmensleistungen mit ihren eigenen Prognosen in Übereinstimmung zu bringen.⁵³ Diese Strategie zielt darauf

⁴² Vgl. Todd et al. (2024, p. 2).

⁴³ Vgl. M. Baker und Wurgler (2007, p. 130).

⁴⁴ Vgl. Zhou (2018, pp. 241f).

⁴⁵ Vgl. Collins et al. (2024, pp. 2 und 8).

⁴⁶ Vgl. Arslan-Ayaydin et al. (2016, pp. 133f).

⁴⁷ Vgl. De Amicis et al. (2021); Blau et al. (2015, p. 206).

⁴⁸ Vgl. Collins et al. (2024, pp. 7f).

⁴⁹ Vgl. Gregory (2021).

⁵⁰ Vgl. Jiang et al. (2019, p. 148).

⁵¹ Vgl. Jiang et al. (2019, p. 128).

⁵² Vgl. Bochkay et al. (2019, p. 56).

⁵³ Vgl. Pouryousof et al. (2023); Ajinkya und Gift (1984, p. 443).

ab, Unterschiede im Informationsstand zu reduzieren und die Glaubwürdigkeit zu stärken, um dadurch die Marktreaktionen optimal zu steuern.⁵⁴ Die Formulierung der Erzählungen gewährt den Investoren Einblick in die Erwartungen der Manager bezüglich künftiger Entwicklungen des Unternehmens. In der wissenschaftlichen Literatur konnte ein positiver Zusammenhang zwischen einer positiven Ausdrucksweise und Leistungskennzahlen, wie beispielsweise Gewinnen und Cashflows, festgestellt werden. Eine übermäßig positive Darstellung der Manager in narrativen Offenlegungen wird durch verschiedene Faktoren limitiert.⁵⁵ Ein wesentlicher Kostenfaktor ist dabei der potenzielle Verlust an Glaubwürdigkeit, die für Investoren von entscheidender Bedeutung ist. Zudem kann Optimismus in finanziellen Offenlegungen als Hinweis auf eine hohe Branchenrentabilität interpretiert werden, was potenzielle Wettbewerber dazu ermutigen könnte, in die Branche einzutreten. Schließlich birgt die Möglichkeit potenzieller Rechtsstreitigkeiten, die durch übermäßig positive Angaben in qualitativen Offenlegungen hervorgerufen werden können, ein bedeutendes Risiko.⁵⁶

Die Betrachtung der Ausdrucksweise von Unternehmensvertretern in Earnings Calls als objektiver Indikator für private Informationen von Managern über zukünftige Unternehmensleistungen vernachlässigt jedoch die Tatsache, dass Manager von einer Vielzahl von Anreizen und Vorurteilen beeinflusst werden können, die die Transparenz der Offenlegung beeinträchtigen können.⁵⁷ In der Konsequenz kann es zu einer irreführenden Wahrnehmung seitens der Anleger kommen.⁵⁸ Der Agency-Theorie zufolge treffen Manager mit Insiderwissen häufig Entscheidungen, die ihre eigenen Interessen mehr berücksichtigen als die der Aktionäre.⁵⁹ Opportunistische Manager tendieren demnach dazu, ihre Schwächen zu verbergen und ihre Erfolge übermäßig hervorzuheben, um positive Berichterstattung zu maximieren und negative Konsequenzen zu minimieren.⁶⁰ Eine negative Stimmung unter Investoren sowie ungünstige Marktreaktionen können Manager in eine Drucksituation bringen, in der persönliche Faktoren wie Vergütung, Karriereentwicklung und Erhalt des eigenen Rufs eine ausschlaggebende Rolle spielen.⁶¹ Beispielsweise legten Arslan-Ayaydin et al. (2016) dar, dass Manager mit aktienbasierter Vergütung eine positivere Tonalität in Gewinnmitteilungen wählen, um über steigende Kurse ihre Vergütung zu erhöhen.⁶² Huang et al. (2014) stellten fest, dass Unternehmen, die im Kontext spezifischer Ereignisse wie beispielsweise der Ausgabe neuer Aktien oder Fusionen eine positive Wahrnehmung bei Investoren anstreben, dazu tendieren, einen optimistischen Tonfall in der entsprechenden Un-

ternehmenskommunikation zu verwenden.⁶³ Es besteht also die Möglichkeit, dass Anleger die in nicht-finanziellen Berichten ausgedrückte Stimmung der Unternehmensführung aufgreifen, auch wenn diese nicht den tatsächlichen Wertgrundlagen des Unternehmens entspricht.⁶⁴ Eine übermäßig positive (negative) Berichterstattung des Managements könnte folglich zu einer spekulativen Überbewertung (Unterbewertung) des Marktes führen.

Darüber hinaus wird von der Verhaltenstheorie angenommen, dass eine gewisse Inkonsistenz im Tonfall von Managern entweder zufällige oder persönliche Ursachen hat.⁶⁵ Davis et al. (2015) demonstrierten, dass Manager-spezifisch optimistische (oder pessimistische) Tendenzen den Sprachstil der finanziellen Offenlegungen über die aktuelle oder zukünftige finanzielle Unternehmensleistung und strategische Anreize der Manager hinaus beeinflussen.⁶⁶ Des Weiteren lässt sich feststellen, dass bestimmte beobachtbare Faktoren, wie frühe berufliche Erfahrungen und ehrenamtliches Engagement, einen Einfluss auf die Art und Weise haben, wie Manager kommunizieren. So konnten De Amicis et al. (2021) und Klevak et al. (2024) geschlechtsspezifische Unterschiede im Tonfall nachweisen.⁶⁷

In einigen Forschungsarbeiten wird folglich die These aufgestellt, dass Nuancen im Tonfall der Bekanntgabe von Unternehmensmeldungen lediglich begrenzte, wenn überhaupt messbare, Auswirkungen auf die langfristige Kursentwicklung haben.⁶⁸ In diesem Zusammenhang wird das in Unternehmensveröffentlichungen zum Ausdruck gebrachte Sentiment von mehreren Autoren als „leeres Gerede“ bezeichnet. Daraus lässt sich die folgende Hypothese ableiten:

H1c: *Die initiale Marktreaktion auf das Sentiment in Earnings Calls erfährt über einen längeren Zeitraum hinweg eine Korrektur*

Vor diesem Hintergrund ist es ebenfalls denkbar, dass die Präsentation sowie die anschließende Fragerunde unterschiedliche Dimensionen des Management-Sentiments widerspiegeln, die jeweils separat die Marktreaktionen beeinflussen können.⁶⁹ Das spezifische Sentiment in dem spontaneren Q&A-Abschnitt bietet möglicherweise detailliertere und authentischere Einblicke in die wahre Unternehmenslage als die strategisch vorformulierte Darstellung in der Präsentation. Da Manager im Q&A auf kritische und unerwartete Fragen reagieren müssen, könnten etwaige Schwächen oder Unsicherheiten deutlicher werden und die persönliche Haltung sowie der emotionale Zustand des Managements unmittelbarer reflektiert werden. Daraus ergeben sich die folgenden zwei Hypothesen:

⁵⁴ Vgl. Rahman (2023).

⁵⁵ Vgl. Arslan-Ayaydin et al. (2016, p. 134).

⁵⁶ Vgl. Rogers et al. (2011, p. 2179).

⁵⁷ Vgl. Arslan-Ayaydin et al. (2016, pp. 133f.); Beyer et al. (2010, pp. 196f).

⁵⁸ Vgl. Pouryousof et al. (2023).

⁵⁹ Vgl. Jensen und Meckling (1976, pp. 308f. and 312f.)

⁶⁰ Vgl. Rahman (2023); Mather et al. (2021, p. 3); Matsumoto et al. (2011, p. 230).

⁶¹ Vgl. R. Luo und Ye (2023, p. 837).

⁶² Vgl. Arslan-Ayaydin et al. (2016, p. 133).

⁶³ Vgl. Huang et al. (2013, p. 1111).

⁶⁴ Vgl. Gupta (2019, p. 17); Jiang et al. (2019, p. 127).

⁶⁵ Vgl. Pouryousof et al. (2023).

⁶⁶ Vgl. Davis et al. (2014, p. 671).

⁶⁷ Vgl. Klevak et al. (2024); De Amicis et al. (2021).

⁶⁸ Vgl. Hennig et al. (2023).

⁶⁹ Vgl. Brockman et al. (2015, pp. 24f).

H2a: *Das Sentiment der Präsentation und das Sentiment im Q&A-Abschnitt hat jeweils einen signifikanten und unabhängigen Einfluss auf die Marktreaktion*

H2b: *Das Sentiment des Q&A-Abschnitts hat einen stärkeren Einfluss auf die Marktreaktion als das Sentiment der Präsentation.*

Neben der Unternehmensführung nehmen auch Analysten eine bedeutsame Rolle in Earnings Calls ein. Sie verfügen über einen umfassenderen Zugang zu Informationsquellen als durchschnittliche Investoren und spielen eine entscheidende Rolle als Vermittler von Unternehmensinformationen.⁷⁰ Im Vergleich zu Managern verfügen Analysten über einen Anreizvorteil, da ihre Interessen in höherem Maße mit jenen der externen Investoren übereinstimmen.⁷¹ Während Manager dazu neigen, Informationen in einem optimistischen Licht darzustellen, nehmen Analysten eine kritischere Haltung ein und legen ein stärkeres Augenmerk auf potenzielle Risiken. Die Ausdrucksweise von Analysten unterliegt jedoch auch bestimmten Einschränkungen. Unternehmen neigen dazu, überwiegend wohlgesonnene Analysten einzuladen.⁷² Zudem lässt sich feststellen, dass Analysten ihren Kommunikationsstil strategisch einsetzen, um eine positive Beziehung zur Unternehmensführung zu etablieren und ihre Informationsgrundlage zu sichern.⁷³ Zur Untersuchung der Unterschiede zwischen Manager- und Analysten-Sentiment werden die Hypothesen formuliert:

H2c: *Das Sentiment der Manager und Analysten innerhalb des Q&A-Abschnitts hat jeweils einen signifikanten und unabhängigen Einfluss auf die Marktreaktion*

H2d: *Das Sentiment der Fragen von Analysten hat einen stärkeren Einfluss auf die Marktreaktion als das Sentiment der Antworten von Managern innerhalb des Q&As*

2.3 Sentimentextraktion mittels textueller Analysen

Die gewonnenen Erkenntnisse veranschaulichen die Komplexität der Informationsübermittlung in Unternehmensoffenlegungen und verdeutlichen die Herausforderungen für Investoren, authentische Signale korrekt zu interpretieren. Vor diesem Hintergrund erweist sich die Sentimentanalyse als ein entscheidendes Instrument, um die psychologischen Einflussfaktoren auf die Märkte zu quantifizieren und zu analysieren.⁷⁴ Es ist daher von entscheidender Bedeutung, zu verstehen, wie Sentiment am besten erfasst werden kann. Zur Ermittlung von Sentiment

im Kontext finanzieller Analysen werden im Wesentlichen drei Ansätze verfolgt: die Beobachtung von Marktdaten, die Durchführung von Umfragen und die Auswertung von Texten. Kritikpunkte an marktbasierenden Ansätzen sind die begrenzte Datenverfügbarkeit, der vergangenheitsbezogene Charakter und Endogenitätsprobleme.⁷⁵ Umfragen bieten zwar wertvolle Einblicke in den Meinungsbildungsprozess der Marktteilnehmer, sind jedoch mit einem hohen Aufwand bei der Durchführung verbunden. Des Weiteren ist die Datenverfügbarkeit hinsichtlich des Umfangs und der Häufigkeit eingeschränkt und die Ergebnisse hängen im hohem Maße vom Erhebungsdesign ab, wobei die Gefahr von Antwortverzerrungen besteht.⁷⁶ Mit Entwicklungen in der Rechenleistung, der Computerlinguistik und der Verfügbarkeit von Daten hat daher die Extraktion von Sentiment aus Texten an Bedeutung gewonnen. Dies resultierte in einer verstärkten Untersuchung linguistischer Signale in der Fachliteratur für Rechnungswesen und Finanzen. Die textuelle Analyse erlaubt die Extraktion sprachlicher Informationen, die für herkömmliche Leser oft nicht offensichtlich sind.⁷⁷ Dabei geht die textbasierte Sentimentanalyse über die bloße Inhaltsbewertung hinaus, indem sie darauf abzielt, die emotionale Haltung oder persönliche Meinung hinter den Texten zu erfassen und zu quantifizieren. Textuelle Auswertungen schneiden bei der Aktienmarktprognose besser ab als markt- und umfragebasierte Methoden.⁷⁸ Dies deutet darauf hin, dass sie in diesem Bereich bislang unzureichend beachtet oder berücksichtigt werden.

Die Wechselwirkung zwischen textuellem und marktteilnehmerbezogenem Sentiment ist dabei komplex und noch nicht vollständig erforscht, einschließlich der Art und Weise, wie Adressaten das textuelle Sentiment deuten.⁷⁹ Zur Untersuchung des Sentiments erfolgt vielfach eine Analyse des „Tons“ von Texten, wobei dieser sowohl vom Inhalt als auch von der Wortwahl abhängt.⁸⁰ Ein positiver Ton kann demnach durch die Hervorhebung positiver Ergebnisse und/oder einer positiven Berichterstattung über die Ergebnisse erreicht werden. Ein gängiges Verfahren zur Bestimmung des Tons ist die Ermittlung der Anzahl positiver und negativer Wörter. Dabei lässt ein überwiegender Anteil positiver Wörter im Vergleich zu negativen auf einen optimistischeren Ton schließen. Der Begriff des textuellen Sentiments ist jedoch nicht allein auf die Dimensionen Positivität und Negativität beschränkt, sondern kann auch auf andere, darüberhinausgehende Dimensionen anwendbar sein, wie etwa stark-schwach und aktiv-passiv.⁸¹ Im weiteren Verlauf dieser Untersuchung wird der Ton eines Textes, gemessen an der relativen Häufigkeit positiver und negativer Wörter, als Indikator für das vorherrschende Sentiment herangezogen.

⁷⁵ Vgl. Bai et al. (2023).

⁷⁶ Vgl. Zhou (2018, pp. 248–250).

⁷⁷ Vgl. Pennebaker et al. (2003, p. 554).

⁷⁸ Vgl. Zhou (2018, p. 250).

⁷⁹ Vgl. Kearney und Liu (2014, p. 172).

⁸⁰ Vgl. Bassyouny et al. (2022); Henry (2008, p. 377).

⁸¹ Vgl. Kearney und Liu (2014, p. 172).

⁷⁰ Vgl. Y. Chen et al. (2023).

⁷¹ Vgl. Brockman et al. (2015, p. 25).

⁷² Vgl. R. Luo und Ye (2023, p. 834); Lee (2015, p. 229)

⁷³ Vgl. Milián et al. (2017, pp. 14f.); Brockman et al. (2015, p. 25).

⁷⁴ Vgl. Zhou (2018, pp. 204–242).

Die Anfänge der Textanalyse in der Finanzforschung demonstrieren das Potenzial der Untersuchung von Marktbebewegungen unter Einbezug von Sentiment aus verschiedenen Textquellen. In einer der ersten Anwendungen der Textanalyse im Jahr 2004 untersuchten Antweiler und Frank den Einfluss von Börsenforen auf den Aktienmarkt.⁸² Sie stellten fest, dass das Nachrichtenvolumen in Börsenforen in der Lage ist, die Volatilität des Marktes vorherzusagen. In Anlehnung an den Ton der Nachrichten in diesen Foren konnten Das und Chen (2007) nachweisen, dass die Entwicklung von Indexrenditen prognostizierbar ist.⁸³ Im selben Jahr zeigte Tetlock, dass das in einer Kolumne des Wall Street Journal ausgedrückte Sentiment die zukünftige Entwicklung der Aktienkurse vorhersagen kann.⁸⁴ In der Folge wurde in der Literatur ein breites Spektrum unterschiedlicher Textsorten analysiert, wie Nachrichtenartikel,⁸⁵ soziale Medien⁸⁶ und Unternehmensmitteilungen.⁸⁷

Insbesondere die Analyse des in Unternehmensveröffentlichungen geäußerten Sentiments erweist sich als hilfreich, um die Bedeutung qualitativer Informationen für die Bewertung einzelner Unternehmen und deren Kapitalmarktpromerance besser zu verstehen.⁸⁸ Obwohl die finanzielle Berichterstattung oft mit Zahlen in Verbindung gebracht wird, ist letztlich Sprache das primäre Kommunikationsmedium für die Vermittlung vergangener und zukünftiger Unternehmensleistungen.⁸⁹ Die unzureichende Regulierung der Tonalität in der Finanzberichterstattung eröffnet der Unternehmensleitung einen größeren Spielraum bei der Hervorhebung bestimmter Aspekte der Unternehmensleistung sowie der Wortwahl zur Beschreibung dieser Leistung.⁹⁰ Es besteht die Möglichkeit, dass zwei Unternehmen, deren Finanzkennzahlen identisch sind, Offenlegungen herausgeben, die eine sehr unterschiedliche Gesamtbotschaft bezüglich der Unternehmensleistung vermitteln. Der Großteil qualitativer Informationen über Unternehmen ist öffentlich zugänglich und in Textform verfügbar.⁹¹ Im Rahmen der Untersuchung von Unternehmensmitteilungen wird insbesondere analysiert, wie optimistische und pessimistische Sprache von Managern in narrativen Berichten eingesetzt wird, um wesentliche Informationen über ein Unternehmen zu kommunizieren.⁹² Dabei wurden Textanalysen verschiedener Arten von Unternehmensveröffentlichungen durchgeführt, darunter Pressemitteilungen,⁹³ Gewinnberichten⁹⁴ und Managementkommentare in den Jahres-

und Quartalsberichten.⁹⁵

Insbesondere im Kontext von Earnings Calls eröffnet die textuelle Analyse die Möglichkeit, die Einschätzungen und Bewertungen der Teilnehmer in einer systematischen Weise zu erfassen und zu evaluieren. Computer können in diesen schriftlichen Aufzeichnungen Nuancen im Tonfall der Unternehmensleitung identifizieren, die von Analysten möglicherweise übersehen werden.⁹⁶ Die Anwendung von Computeralgorithmen zur Sentimentanalyse kann demzufolge zu einer präziseren und umfassenderen Erfassung der kommunizierten Informationen führen. Da es keine gesetzliche Verpflichtung gibt, schriftliche Transkripte von Earnings Calls bereitzustellen, sind nicht alle Telefonkonferenzen dokumentiert.⁹⁷ Dennoch werden sie von einer steigenden Anzahl an Unternehmen oder Drittanbietern transkribiert. Das Fehlen solcher Dokumente kann insbesondere für Einzel- und nicht-professionelle Investoren von Nachteil sein.⁹⁸ Obwohl die meisten Unternehmen ihre Earnings Calls zeitnah nach Veröffentlichung der Ergebnisse über Webcasts oder Einwahldienste live zugänglich machen, sind die enthaltenden Informationen oft auch über den Termin hinaus relevant. Da mehrere Konferenzen gleichzeitig stattfinden können, ist es für Anleger oft schwierig, alle in Echtzeit zu verfolgen. Schriftliche Transkripte erlauben zudem eine erleichterte Verarbeitung der Informationen, da sie eine wiederholbare und flexible Analyse ermöglichen, während die Verarbeitung von Informationen in Live-Calls in Echtzeit erfolgen muss und stimmliche sowie visuelle Ablenkungen zusätzliche kognitive Belastungen mit sich bringen.⁹⁹ Vor diesem Hintergrund erweisen sich Transkripte als effiziente Alternative, die für zahlreiche Anleger einen wesentlichen Bestandteil ihrer Anlageforschung darstellen.

Die Literatur zum Einfluss von Sentiment in Earnings-Call-Transkripten auf die Kapitalmärkte bietet vielfältige Einblicke in die Dynamik zwischen Unternehmenskommunikation und Marktreaktionen. In einer Studie von Price et al. (2012) wurde erstmals der Zusammenhang zwischen dem Sentiment in diesen Telefonkonferenzen und Kapitalmarktreaktionen untersucht.¹⁰⁰ Die Ergebnisse zeigen, dass das Sentiment in Earnings Calls einen signifikanten Einfluss auf verschiedene Finanzmarktgrößen ausübt. Dazu zählen abnormale Renditen sowohl während des Gewinnankündigungsfensters als auch über einen längeren Zeitraum hinweg sowie das Handelsvolumen. Borochin et al. (2017) ergänzten diese Erkenntnisse, indem sie feststellten, dass pessimistische Stimmungen während Earnings Calls zu größerer Markttunsicherheit führen als optimistische Stimmungen.¹⁰¹ Insbesondere ist der starke Einfluss einer negativen Analystenstimmung während des Q&A-Abschnitts hervorzuheben. Dies

⁸² Vgl. Antweiler und Frank (2004, p. 1292).

⁸³ Vgl. Das und Chen (2007, p. 1386).

⁸⁴ Vgl. Tetlock (2007, p. 1166).

⁸⁵ Vgl. Bai et al. (2023, p. 8); Tetlock (2007, p. 1166).

⁸⁶ Vgl. Agrawal et al. (2018, p. 86); Renault (2017, p. 37).

⁸⁷ Vgl. Jiang et al. (2019, p. 148); Loughran und McDonald (2011, p. 61).

⁸⁸ Vgl. Kearney und Liu (2014, p. 173).

⁸⁹ Vgl. Hales et al. (2011, p. 224).

⁹⁰ Vgl. Henry und Leone (2016, p. 153).

⁹¹ Vgl. Kearney und Liu (2014, p. 181).

⁹² Vgl. Bassyouny et al. (2022).

⁹³ Vgl. Arslan-Ayaydin et al. (2016, p. 133); Huang et al. (2013, p. 1111).

⁹⁴ Vgl. DeBoskey et al. (2018, p. 658); Y. Luo und Zhou (2017, p. 454); Tama-Sweet (2014, p. 276).

⁹⁵ Vgl. Yan et al. (2023, p. 2); D'Augusta und DeAngelis (2019, p. 133).

⁹⁶ Vgl. Loughran und McDonald (2016, p. 1188).

⁹⁷ Vgl. Elliott et al. (2023, pp. 9 und 12).

⁹⁸ Vgl. Altman (2019).

⁹⁹ Vgl. Elliott et al. (2023, pp. 9 und 12).

¹⁰⁰ Vgl. Price et al. (2012, p. 1006).

¹⁰¹ Vgl. Borochin et al. (2018, p. 76).

führen die Autoren darauf zurück, dass Analysten als vertrauenswürdiger Informationsquellen gelten als Unternehmensvertreter. Fu et al. (2019) stellten fest, dass ein höherer Optimismus in Earnings Calls das Risiko eines Aktienkursrückgangs signifikant reduzieren kann.¹⁰² Auch sie identifizierten den Q&A-Teil der Calls als besonders informativ, was darauf hindeutet, dass die Interaktion zwischen Analysten und Managern entscheidend für die besonderen Marktreaktionen auf Earnings Calls ist.

In der wissenschaftlichen Literatur wird vielfach die Auffassung vertreten, dass das Sentiment der Analysten im Q&A-Abschnitt von Earnings Calls eine größere Relevanz besitzt als das der Manager. Brockman et al. (2015) zeigten, dass Manager in Earnings Calls eine tendenziell optimistischere Haltung als andere Teilnehmer einnehmen.¹⁰³ Investoren vertrauen in erster Linie dem Ton von Analysten, die ähnliche Anreize wie sie haben, obwohl Manager über einen erheblichen Informationsvorteil verfügen. Ergänzend stellten X. Chen und Tian (2018) fest, dass Analysten Manager dazu bringen können, ihren anfänglichen Optimismus aufzugeben, und dass Aktienkurse stärker auf den Ton der Analysten als auf den der Manager reagieren.¹⁰⁴ Im Gegensatz dazu wiesen Milian et al. (2017) darauf hin, dass der Ton des gesamten Earnings Calls Aktienrenditen vorhersagen kann, nicht jedoch der der Analysten.¹⁰⁵ Auch Yang et al. (2023) zeigten, dass die Aussagen der Manager im Q&A-Teil des Earnings Calls für den Markt bedeutsamer sind als die der Analysten.¹⁰⁶ Diese widersprüchlichen Erkenntnisse verdeutlichen, dass der Einfluss von Sentiment in verschiedenen Segmenten von Earnings Calls auf Investoren noch nicht vollständig erforscht ist. Die unterschiedlichen Ergebnisse könnten insbesondere auf Variationen in der Informationslage der Anleger sowie auf methodische Unterschiede in den Studien zurückzuführen sein. Weitere Forschung ist erforderlich, um diese Diskrepanzen zu klären und ein umfassenderes Verständnis dafür zu entwickeln, wie und warum Sentiment in verschiedenen Segmenten von Earnings Calls unterschiedliche Auswirkungen auf den Markt hat.

2.4 Lexikonbasierte Textauswertung

Die in der Literaturübersicht dargestellten Beobachtungen verdeutlichen, dass aufgrund der inhärenten Ungenauigkeiten dieses Vorhabens die Methodik von zentraler Bedeutung ist, mit der qualitative Stimmungen, Einstellungen und Emotionen in Texten quantifiziert und in ein Sentimentmaß umgewandelt werden.¹⁰⁷ Während quantitative Forschung auf etablierten Normen und klar definierten Maßeinheiten basiert, erfordert die Auswertung von Texten aufgrund der Mehrdeutigkeit sprachlicher Zeichen eine kontextabhängige

Interpretation. In der frühen Kapitalmarktforschung wurden subjektive Bewertungen des textuellen Tons durch menschliche Codierer vorgenommen.¹⁰⁸ Heute werden nahezu ausschließlich computergestützte Programme zur Verarbeitung natürlicher Sprache eingesetzt, um Indikatoren für das Sentiment der Offenlegung zu ermitteln. Diese Verfahren lassen sich der lexikonbasierten Analyse oder dem maschinellen Lernen zuordnen.

Der prominenteste Ansatz zur Ermittlung der Stimmung eines Textes besteht darin, den Text in seine Wörter zu zerlegen und diese anschließend mithilfe eines Lexikons zu kategorisieren.¹⁰⁹ Für diese lexikonbasierten Textanalysen wurden vordefinierte Wörterbücher entwickelt, die Wörtern Attribute wie „positiv“ oder „negativ“ zuweisen. Diese Methodik lässt sich dem sogenannten „Bag of Words“-Ansatz zuordnen. Dabei wird die Häufigkeit von Wörtern innerhalb eines Textdokumentes untersucht, ohne die Reihenfolge und den unmittelbaren Zusammenhang der Wörter zu berücksichtigen. Der lexikonbasierte Ansatz kann demnach Schwierigkeiten aufweisen, den Kontext und die Mehrdeutigkeit von Wörtern adäquat zu erfassen.¹¹⁰ Zudem besteht die Möglichkeit, dass Lexika veraltet sind und Änderungen in der Sprache nicht angemessen widerspiegeln.¹¹¹

Technologische Fortschritte haben in den letzten Jahren die Möglichkeiten, maschinelles Lernen in der Finanzanalyse einzusetzen, erheblich erweitert und bieten somit eine wertvolle Ergänzung zu den traditionellen, lexikonbasierten Ansätzen.¹¹² Der Begriff des maschinellen Lernens bezeichnet die Fähigkeit von Computern, aus Daten Muster zu erkennen und Vorhersagen zu treffen.¹¹³ Die Anwendung von Algorithmen und statistischen Modellen erlaubt die Übertragung der gewonnenen Erkenntnisse auf andere Daten. Zu den Vorteilen maschinellen Lernens im Kontext textueller Analysen gehören demnach die Fähigkeit, sich an Veränderungen anzupassen sowie eine potenzielle höhere Genauigkeit bei spezifischen Aufgaben. Ein Beispiel für den Einsatz von maschinellem Lernen liefert die Studie von Frankel et al. (2021).¹¹⁴ In dieser wurden verschiedene Algorithmen, darunter neuronale Netze, eingesetzt, um Klassifikationsmodelle zu trainieren, die in der Lage sind, das Sentiment in Jahresberichten und Transkripten von Earnings Calls zu erkennen. Die Resultate demonstrieren, dass maschinelle Lernmodelle eine höhere Präzision aufweisen als konventionelle, auf Lexika basierende Ansätze. Die von Renault (2020) durchgeführte Analyse verdeutlicht jedoch, dass komplexe Algorithmen nicht zwangsläufig präzisere Ergebnisse liefern als lexikonbasierte Methoden.¹¹⁵

Vor diesem Hintergrund erweist sich der lexikalische Ansatz, der im Rahmen der empirischen Untersuchung An-

¹⁰⁸Vgl. Henry und Leone (2016, p. 156).

¹⁰⁹Vgl. Loughran und McDonald (2016, pp. 1198–1200).

¹¹⁰Vgl. Loughran und McDonald (2016, p. 1226).

¹¹¹Vgl. Frankel et al. (2021, p. 5516).

¹¹²Vgl. Todd et al. (2024, p. 4).

¹¹³Vgl. Frankel et al. (2021, pp. 5517f).

¹¹⁴Vgl. Frankel et al. (2021, pp. 5514f).

¹¹⁵Vgl. Renault (2017, p. 31).

¹⁰²Vgl. Fu et al. (2019, p. 644).

¹⁰³Vgl. Brockman et al. (2015, p. 40).

¹⁰⁴Vgl. X. Chen und Tian (2018, p. 1348).

¹⁰⁵Vgl. Milian et al. (2017, p. 73).

¹⁰⁶Vgl. Yang et al. (2022, p. 2177).

¹⁰⁷Vgl. Loughran und McDonald (2016, pp. 1189–1191).

wendung findet, weiterhin als eine bewährte Methode, die gegenüber maschinellen Lernansätzen eine Reihe von Vorteilen aufweist. Die Verwendung einer vordefinierten Wortliste im Rahmen eines lexikonbasierten Ansatzes gewährleistet eine einfache Implementierung sowie Reproduzierbarkeit der Forschungsergebnisse.¹¹⁶ Im Gegensatz zu maschinellen Lernverfahren benötigen lexikalische Analysen keine umfangreichen Trainingsdaten und ermöglichen eine schnelle Implementierung. Nach Festlegung des zugrunde liegenden Wörterbuchs sind die Ergebnisse objektiv und robust gegen Überanpassung. Maschinelle Lernverfahren werden häufig als Blackbox wahrgenommen, während der lexikonbasierte Ansatz eine transparente und nachvollziehbare Analyse der Ergebnisse ermöglicht. Das Verfahren eignet sich insbesondere für die Analyse großer Datenmengen, da die computergestützte Häufigkeitszählung der Wörter vergleichsweise geringe Rechenressourcen erfordert. Lexika, die gezielt auf die Eigenschaften des vorliegenden Textes abgestimmt sind, liefern bei der Sentiment-Extraktion Ergebnisse, die häufig mit denen differenzierterer Instrumente wie maschinellen Lernens vergleichbar sind.¹¹⁷ Die Auswahl der Wortliste, die als Grundlage für die Klassifizierung von Wörtern dient, erweist sich bei einem lexikonbasierten Ansatz als kritischer Faktor.¹¹⁸ Zur Prüfung der Effekte des ausgewählten Lexikons auf das ermittelte Sentiment wird folgende Hypothese aufgestellt:

H3a: *Der Zusammenhang zwischen Sentiment in Earnings Calls und der Marktreaktion variiert je nach dem verwendeten Sentiment-Lexikon*

Zu Beginn der textbasierten Analysen in den Bereichen Rechnungswesen und Finanzen wurde auf allgemeine Wortlisten wie das Harvard-GI-Lexikon zurückgegriffen, da dieses früh verfügbar war und zu diesem Zeitpunkt noch keine spezielleren Listen für diese Bereiche existierten.¹¹⁹ Stone und Hunt entwickelten im Jahr 1961 das Lexikon für das General Inquirer System.¹²⁰ Dieses computergestützte Analysewerkzeug verwendet vordefinierte Wortlisten, um psychologische und soziale Phänomene in Texten zu erfassen und zu quantifizieren. Das Harvard-GI-Lexikon enthält 2.005 negative und 1.637 positive Wörter, wobei auch verschiedene Wortformen berücksichtigt werden.¹²¹

Die Anwendung generischer Lexika zur Erfassung finanziellen Sentiments kann zu Ungenauigkeiten führen, da sie nicht speziell für den Kontext der Finanzwirtschaft entwickelt wurden.¹²² Um die spezifischen Nuancen des Finanzwesens angemessen erfassen zu können, ist es erforderlich, die Wörterbücher kontextspezifisch anzupassen oder zu erweitern. Loughran und McDonald (2011) berichteten, dass

nahezu 75 % der als Harvard-GI-negativ klassifizierten Wörter in Finanzdokumenten keine negative Bedeutung aufweisen.¹²³ So wird beispielsweise das Wort „beat“ (schlagen) im GI-Wörterbuch als negativ eingestuft, obwohl es in Gewinnmitteilungen häufig verwendet wird, um das Übertreffen von Erwartungen zu beschreiben.¹²⁴ Auch im Hinblick auf die Analyse von Earnings Calls hat sich die Verwendung kontextspezifischer Wörterbücher als überlegen erwiesen.¹²⁵ In ihrer Studie aus dem Jahr 2012 konnten Price et al. nachweisen, dass das finanzspezifische Henry-Wörterbuch im Vergleich zum generischen Harvard-GI eine präzisere Methode zur Ermittlung eines Tons aus Earnings Calls darstellt, der Rückschlüsse auf Aktienrenditen zulässt.¹²⁶ Auch Davis et al. (2015) wiesen darauf hin, dass die kontextspezifischen Henry- und LM-Wörterbücher ein generisches Lexikon bei der Bestimmung des Tons in Earnings Calls übertreffen.¹²⁷

Die erste speziell für finanzielle Texte entwickelte Wortliste wurde im Jahr 2008 von Henry erstellt.¹²⁸ Sie basiert auf der Analyse von Gewinnmitteilungen der Telekommunikations- und Computerdienstleistungsbranche der Jahre 1998 bis 2002. Im Vergleich zu generischen Wörterbüchern wie GI umfasst das Henry-Lexikon nur eine begrenzte Anzahl an Wörtern, nämlich 105 positive und 85 negative.¹²⁹ In der Konsequenz findet eine Vielzahl von im Bereich der Geschäftskommunikation häufig verwendeten Wörtern keine Berücksichtigung, was eine Beeinträchtigung der Erfassung der für Finanztexte relevanten Attribute und Töne zur Folge haben könnte. Um Sentiment in Geschäftsberichten und Finanzdokumenten besser bewerten zu können, erstellten Loughran und McDonald daher das LM-Wörterbuch.¹³⁰ Das Lexikon umfasst die Kategorien negativ, positiv, unsicher, streitlustig, stark modal und schwach modal. Die Zuordnung der Wörter erfolgte durch die Analyse von US-Jahresberichten aus den Jahren 1994 bis 2008, mit dem Ziel, eine umfassende Liste von Wörtern zu erstellen, die nicht gezielt von Unternehmen vermieden werden können. Das LM-Lexikon wird in einem zweijährigen Rhythmus aktualisiert und umfasst zum Zeitpunkt der Erstellung dieser Arbeit 2.355 negative und 354 positive Wörter. Im Vergleich zu anderen, in der Rechnungswesen- und Finanzliteratur verbreiteten Wortwörterbüchern weist es demnach zwei wesentliche Vorteile auf: Einerseits wurde es speziell für die Finanzkommunikation erstellt und basiert auf der wahrscheinlichsten Interpretation eines Wortes im Geschäftskontext, andererseits zeichnet es sich durch eine breite Abdeckung von Wörtern aus. Daher wird die folgende Hypothese getestet:

¹¹⁶Vgl. Frankel et al. (2021, pp. 5516f.); Loughran und McDonald (2016, p. 1200).

¹¹⁷Vgl. Renault (2017, p. 37); Henry und Leone (2016, p. 174).

¹¹⁸Vgl. Loughran und McDonald (2011, p. 36).

¹¹⁹Vgl. Loughran und McDonald (2016, p. 1201).

¹²⁰Vgl. Stone und Hunt (1963, pp. 241f).

¹²¹Vgl. Bassyouny et al. (2022); Loughran und McDonald (2016, p. 1201).

¹²²Vgl. Loughran und McDonald (2016, p. 1226).

¹²³Vgl. Loughran und McDonald (2011, p. 49).

¹²⁴Vgl. Henry und Leone (2016, p. 157).

¹²⁵Vgl. Todd et al. (2024, p. 4).

¹²⁶Vgl. Price et al. (2012, p. 1006).

¹²⁷Vgl. Davis et al. (2014, p. 657).

¹²⁸Vgl. Henry (2008, p. 384).

¹²⁹Vgl. Loughran und McDonald (2016, p. 1201).

¹³⁰Vgl. Loughran und McDonald (2016, p. 1204).

H3b: *Das Sentiment in Earnings Calls, das mit dem LM-Lexikon ermittelt wird, zeigt einen ausgeprägteren Zusammenhang zur Marktreaktion als das Sentiment, das mit den Lexika von Henry oder Harvard GI ermittelt wird*

3 Methodik zur Analyse von Earnings Calls und Marktreaktionen

3.1 Datenquellen und Beschreibung der Stichprobe

Im Rahmen der empirischen Untersuchung werden Earnings Calls von Unternehmen aus dem S&P Composite 1500 Index betrachtet, der sowohl Unternehmen des S&P 500 als auch mittelgroße und kleinere Unternehmen des S&P MidCap 400 und S&P SmallCap 600 umfasst.¹³¹ Der Index beinhaltet insgesamt 1.506 Aktien, darunter sechs Unternehmen mit jeweils zwei Aktienklassen, und repräsentiert damit rund 90 % der Marktkapitalisierung des US-amerikanischen Aktienmarktes. Aufgrund dieser breiten Abdeckung kann er als maßgeblicher Indikator für den amerikanischen Aktienmarkt angesehen werden. Die Analyse von Unternehmen des S&P Composite 1500 ermöglicht die Untersuchung der unterschiedlichen Marktreaktionen auf Earnings Calls einer breiteren Vielfalt von Unternehmen.

Die Transkripte der Earnings Calls sowie alle weiteren Daten werden von LSEG Workspace bezogen. Um eine diversifizierte Stichprobe mit unterschiedlichen Rendite-Risiko-Profilen zu erhalten, werden für alle Unternehmen des S&P Composite 1500 Quartile für das Buchwert-zu-Markt-Verhältnis (BM-Verhältnis) und die Marktkapitalisierung gebildet. Die Marktkapitalisierung ermöglicht die Berücksichtigung von Unternehmen unterschiedlicher Größe, während das BM-Verhältnis auf mögliche Wachstumserwartungen oder Unterbewertungen hinweisen kann. Durch die Einteilung in Quartile entstehen 16 spezifische Portfolios, die Kombinationen aus Unternehmensgröße und Bewertungsverhältnis darstellen und eine breite Bandbreite von Unternehmen repräsentieren. Aus diesen wird eine Zufallsstichprobe von jeweils 15 Unternehmen gezogen. Das Portfolio 1.1, das die Unternehmen mit der niedrigsten Marktkapitalisierung und dem niedrigsten BM-Verhältnis im S&P Composite 1500 umfasst, beinhaltet jedoch lediglich zwölf Unternehmen. Daher umfasst die Gesamtstichprobe 237 anstelle der ursprünglich angenommenen 240 Unternehmen.

Der Betrachtungszeitraum erstreckt sich vom 01.01.2019 bis zum 15.03.2024. In diesem Zeitraum veröffentlichten die meisten Unternehmen ihre Earnings Calls für die Geschäftsjahre 2019 bis 2023. Da von einem Unternehmen der Stichprobe keine Earnings-Call-Transkripte vorliegen, können nur 236 Unternehmen berücksichtigt werden. Im betrachteten Zeitraum hat jedes Unternehmen durchschnittlich 20,42 Earnings Calls durchgeführt, was etwa fünf Berichtsjahren entspricht und zu einer Gesamtzahl von 4.820 analysierten Earnings Calls führt. Durchschnittlich vergingen zwischen Ende

des jeweiligen Quartals und der Durchführung der Earnings Calls 34 Tage.

Die Datenverarbeitung erfolgt unter Zuhilfenahme der Statistiksoftware R sowie der darin integrierten Pakete. In einem ersten Schritt werden die Transkripte der Earnings Calls eingelesen und aufbereitet. Dabei steht die Schaffung einer einheitlichen und strukturierten Datenbasis für die nachfolgende Analyse im Vordergrund. Die Identifikation von Unternehmensnamen, Datum und Quartal erfolgt mittels spezifischer Mustererkennungstechniken unter Verwendung der Funktion „str_extract“ aus dem R-Paket „stringr“. Nach der Extraktion der relevanten Informationen werden verschiedene Datenbereinigungsverfahren durchgeführt. In diesem Zusammenhang werden unterschiedliche Bezeichnungen desselben Unternehmens in den Transkripten der verschiedenen Quartale vereinheitlicht, um eine konsistente Analyse und Kompatibilität mit anderen Daten zu gewährleisten. Darüber hinaus werden die Definitionen sowie der Disclaimer am Ende der Transkripte entfernt, da diese keinen Inhalt der ursprünglichen Telefonkonferenz wiedergeben.

Im Anschluss wird neben der Betrachtung des gesamten Earnings Calls eine Aufteilung der Transkripte in verschiedene Teile vorgenommen, um eine differenzierte Analyse des Sentiments und seiner Auswirkungen auf die Kursreaktionen zu ermöglichen. In einem ersten Schritt erfolgt eine Trennung der beiden Hauptkomponenten des Calls. Hierbei handelt es sich zum einen um die vorbereitete Präsentation und zum anderen um den anschließenden, eher spontanen Q&A-Abschnitt. Der Q&A-Abschnitt wird, wie nachfolgend beschrieben, weiter in Äußerungen von Unternehmensvertretern und Analysten untergliedert: Zu Beginn jedes Transkripts werden alle Teilnehmer mit ihren jeweiligen Funktionen aufgelistet, wobei den Äußerungen der Name des Sprechers vorangestellt ist. Um eine einheitliche Analyse zu gewährleisten, werden die Teilnehmer in zwei Kategorien eingeteilt: Unternehmensvertreter und sonstige Teilnehmer, wie Analysten. Im Rahmen der Datenaufbereitung werden die Namen im gesamten Transkript durch spezifische Platzhalter ersetzt, die den jeweiligen Teilnehmertyp kennzeichnen. Anschließend wird vor jedem dieser Platzhalter ein einheitliches Trennzeichen eingefügt, sodass die Textabschnitte effizient mithilfe der „strsplit“-Funktion segmentiert und den Funktionen der Sprecher (Manager oder Analysten) zugeordnet werden können.

Die textuelle Analyse von Earnings-Call-Transkripten erfordert eine einheitliche Struktur, um die automatische Verarbeitung, Extraktion und Zuordnung von Informationen, insbesondere in Bezug auf verschiedene Bestandteile und Personen, zu erleichtern. Von den ursprünglich 4.820 betrachteten Calls können ca. 6 % (inklusive aller Calls zweier Unternehmen) aufgrund struktureller Abweichungen nicht für die weitere Analyse verwendet werden. Nach Ausschluss der fehlerhaften Transkripte verbleiben 4.548 Earnings-Call-Transkripte, die eine konsistente und valide Datenbasis für die weiterführende Analyse bieten.

¹³¹Vgl. S&P Global (2024).

3.2 Sentimentermittlung in Earnings Calls

Vor der Analyse des in den Transkripten vermittelten Sentiments ist eine Bereinigung und Vorbereitung der Texte erforderlich. In Anlehnung an die Arbeit von Allee und Deangelis (2015) erfolgt in der vorliegenden Arbeit eine spezifische Anpassung der Textabschnitte für die Analyse von Earnings Calls.¹³² In diesem Zusammenhang wird das Wort „Question“ entfernt, da es in Telefonkonferenzen häufig dazu verwendet wird, Bezug auf die Fragen der Analysten zu nehmen. Des Weiteren wird das Wort „good“ nicht zwangsläufig als positiv kategorisiert, insbesondere wenn es in Begrüßungsformeln wie „good morning“ verwendet wird, die eher konventionellen Höflichkeitsfloskeln eines Dialogs entsprechen. Ebenso wird auch der Ausdruck „Thank you“ ausgeschlossen, der in vielen Fällen dazu dient, sich für die Beantwortung einer Frage zu bedanken.

In einem nächsten Schritt werden Vereinfachungsoperationen an den Textabschnitten durchgeführt, die darauf abzielen, das Datenvolumen und die Komplexität zu reduzieren, ohne dabei die Informationsdichte zu beeinträchtigen.¹³³ Eine Möglichkeit, einen Text zu vereinfachen, besteht darin, Stoppwörter wie Präpositionen, Artikel, Pronomen und Konjunktionen zu eliminieren.¹³⁴ Im Gegensatz zu den „lexikalischen Wörtern“ (Substantive, Adjektive, Verben) enthalten diese nur wenig Information, machen jedoch etwa 20–30 % der Wörter eines Standardtextes aus. Die Umsetzung erfolgt in R mittels der Pakete „NLP“ und „tm“. In einem weiteren Schritt werden die zur Zuweisung der Textabschnitte zu Teilnehmerrollen verwendeten Platzhalter sowie Zahlen, Sonderzeichen und Interpunktionen entfernt. Im Anschluss erfolgt die Umwandlung aller Wörter in Kleinschreibung.

Im Rahmen einer lexikonbasierten Analyse stellt die Lemmatisierung einen weiteren wesentlichen Schritt dar.¹³⁵ Dabei werden Wörter auf ihren jeweiligen Wortstamm zurückgeführt. Dies ist insbesondere von Bedeutung, um sicherzustellen, dass alle relevanten Wörter durch das Lexikon erfasst werden, da vordefinierte Lexika häufig lediglich die Grundform der Wörter enthalten. Im Vergleich zum Stemming, bei dem das Präfix und das Suffix der Wörter entfernt werden, stellt die Lemmatisierung ein komplexeres und zeitintensiveres Verfahren dar. Sie zeichnet sich jedoch durch eine höhere Präzision sowie geringere Fehleranfälligkeit aus. Ein Beispiel für die Auswirkungen unterschiedlicher Verfahren ist die Rückführung des Wortes „better“ durch Stemming auf den Wortstamm „bett“. Im Gegensatz dazu identifiziert die Lemmatisierung „good“ als Wortstamm. Die Funktion „lemmatize_strings“ des Pakets „textstem“ automatisiert die Lemmatisierung der Wörter in R. In Tabelle 1 wird exemplarisch die Bereinigung und Lemmatisierung eines Abschnitts aus einem Earnings Call dargestellt, in welchem ein Manager auf die Frage eines Analysten antwortet.

Tabelle 1: Bereinigen und lemmatisieren der Textabschnitte

Originaler Earnings Call Abschnitt	Bereinigter und lemmatisierter Abschnitt
<p>Yes. We have restructured to some extent, inside the organization to focus much more on customer segments, enterprise, commercial and hyperscaler and putting specialized teams focused on driving business process to fit those 3 models in a very coordinated way across the company. And based on that, we've moved people and moved roles to different areas. And that also meant that there were some roles that were being eliminated, and that's what really caused the restructuring. And no, we don't – I mean, that was limited really to the reorganization that we put in place.¹³⁶</p>	<p>yes restructure extent inside organization focus much customer segment enterprise commercial hyperscaler put specialize team focus drive business process fit model coordinate way across company base move people move role different area also mean role eliminate really cause restructure mean limit really reorganization put place</p>

Im Anschluss erfolgt die Ermittlung der Anzahl positiver und negativer Wörter für jeden Textabschnitt unter Zuhilfenahme der Lexika Harvard-GI (GI), Henry (HE) sowie Lougbran und McDonald (LM). Dazu wird ein Abgleich zu den Positiv-/Negativ-Listen aus den drei verschiedenen Wörterbüchern mit der Funktion „str_count“ aus dem Paket „stringr“ durchgeführt. Die in Tabelle 2 dargestellten Ergebnisse zeigen die Anzahl positiver (+) und negativer (-) Wörter je Wörterbuch, ermittelt aus dem exemplarischen Ausschnitt eines Earnings Calls.

¹³²Vgl. Allee und Deangelis (2015, p. 10).

¹³³Vgl. Tuffery (2023, p. 14).

¹³⁴Vgl. Tuffery (2023, pp. 130f).

¹³⁵Vgl. Tuffery (2023, pp. 124–126).

¹³⁶Giancarlo (2024).

Tabelle 2: Quantifizierung der in den Textabschnitten enthaltenden Stimmung

Bereinigter und lemmatisierter Abschnitt	GI +	GI -	HE +	HE -	LM +	LM -
<i>yes restructure extent inside organization focus much customer segment enterprise commercial hyperscaler put specialize team focus drive business process fit model coordinate way across company base move people move role different area also mean role eliminate really cause restructure mean limit really reorganization put place</i>	9	9	1	0	0	2

In einem weiteren Schritt erfolgt die Transformation der Anzahl positiver ($N_{positiv}$) und negativer ($N_{negativ}$) Wörter je Earnings Call eines Unternehmens i zu dem Zeitpunkt t und Lexikon l in ein relatives Maß für das Sentiment:

$$Sentiment_{i,t,l} = \frac{N_{positiv,i,t,l} - N_{negativ,i,t,l}}{N_{positiv,i,t,l} + N_{negativ,i,t,l}}. \quad (1)$$

Die verwendete Maßeinheit für die Sentiment-Messung ist auf das Intervall $[-1, 1]$ beschränkt. Der Wert -1 steht dabei für eine pessimistische und der Wert 1 für eine optimistische Sprache. Da die automatische Trennung verschiedener Teile des Earnings Calls (etwa in Präsentation und Q&A) mitunter zu Ungenauigkeiten führt, werden Textabschnitte, die eine Mindest- oder Höchstanzahl an Wörtern nicht erfüllen, bei der Ermittlung des Sentiments nicht berücksichtigt.

3.3 Ereignisstudie zur Untersuchung der Marktreaktionen

Die Frequenz, in der Unternehmen Berichte veröffentlichten – typischerweise in Quartal- oder Jahresschritten – begrenzt die Eignung von aus Unternehmensveröffentlichungen ermitteltem Sentiment als Input für Zeitreihenmodelle.¹³⁷ Dennoch eignen sich diese Daten für Querschnittsanalysen und Ereignisstudien, in denen die Auswirkungen der Stimmung auf die Kapitalmarktreaktion rund um Unternehmensveröffentlichungen untersucht werden können. Die Ereignisstudien-Methodik, wie sie heute in der akademischen Forschung Anwendung findet, wurde erstmals 1969 von Fama et al. als Instrument zur Ermittlung ereignisinduzierter Renditen eingeführt.¹³⁸ Sie erlaubt die Quantifizierung des Einflusses eines Ereignisses auf die Rendite einer Aktie innerhalb eines definierten Zeitraums, dem sogenannten „Ereignisfenster“. Folglich kann mit ihrer Hilfe die Marktbewertung des Informationsgehalts von Offenlegungen sowie der allgemeinen Qualität der Finanzberichterstattung erfasst und analysiert werden.¹³⁹

Ereignisstudien basieren auf der Prämisse einer mittelstarken Markteffizienz, bei der Investoren auf die Veröffentlichung oder das Bekanntwerden neuer Informationen (das

Ereignis) reagieren.¹⁴⁰ Eine durch ein Ereignis ausgelöste Änderung des Aktienkurses erfolgt, wenn die zugrunde liegende Information einen positiven oder negativen Informationswert besitzt und für die Marktteilnehmer neu und unerwartet ist.¹⁴¹ Im Vergleich zu älteren Methoden berücksichtigt die moderne Ereignisstudien-Methodik nicht einen Vorher-Nachher-Kursvergleich, sondern die hypothetische Kursentwicklung des Unternehmens, falls das Ereignis nicht eingetreten wäre.¹⁴² Die Differenz zwischen der beobachteten Kursveränderung $R_{i,t}$ und der hypothetisch erwarteten Rendite $E(R_{i,t})$ wird als abnormale Rendite $AR_{i,t}$ bezeichnet und dient als Indikator für die Kapitalmarktbeurteilung eines Ereignisses:

$$AR_{i,t} = R_{i,t} - E(R_{i,t}). \quad (2)$$

In der wissenschaftlichen Literatur finden sich zahlreiche Methoden zur Bestimmung der erwarteten Rendite $E(R_{i,t})$.¹⁴³ Die resultierenden Referenzrenditen können dabei erheblich variieren. Das von Sharpe (1963) auf Basis der Markowitz-Portfolio-Theorie entwickelte Marktmodell stellt eine Variante des Einfaktormodells dar und gilt als nahezu universell einsetzbares Instrument zur Bestimmung erwarteter Renditen in kurzfristigen Eventstudien.¹⁴⁴ Dieses Modell erweist sich insbesondere dann als geeignet, wenn das untersuchte Ereignis, wie auch Earnings Calls, in erster Linie unternehmensspezifisch ist und keine wesentlichen Auswirkungen auf den Gesamtmarkt oder systematische Risiken aufweist.¹⁴⁵ Die Marktrendite $R_{m,t}$ stellt dabei die wesentliche Einflussgröße auf die Rendite des untersuchten Wertpapiers dar.¹⁴⁶ Im Rahmen dieser Arbeit wird die Rendite des S&P Composite 1500 als relevante Marktrendite $R_{m,t}$ definiert. Die Formel zur Bestimmung der erwarteten Rendite lautet:

$$E(R_{i,t}) = \hat{\alpha}_i + \hat{\beta}_i * R_{m,t}, \quad (3)$$

¹⁴⁰Vgl. Gehrke (2022, p. 279).

¹⁴¹Vgl. Holler (2016, p. 23).

¹⁴²Vgl. Gerpott (2009, p. 206).

¹⁴³Vgl. Gehrke (2022, p. 279).

¹⁴⁴Vgl. Henry und Leone (2016, p. 277); Holler (2016, pp. 82f).

¹⁴⁵Vgl. Müller (2023).

¹⁴⁶Vgl. Gehrke (2022, p. 280).

¹³⁷Vgl. Kearney und Liu (2014, p. 174).

¹³⁸Vgl. Holler (2016, p. 11).

¹³⁹Vgl. Henry (2008, p. 283).

wobei die Koeffizienten $\hat{\alpha}_i$ und $\hat{\beta}_i$ eines Unternehmens i für jeden Zeitpunkt t durch eine lineare Regression aus historischen Unternehmens- und Marktrenditen ermittelt werden. Um eine Verzerrung durch das Ereignis selbst zu vermeiden, muss die Schätzung für die Regression des Marktmodells dabei außerhalb des Ereigniszeitraums liegen. In der vorliegenden Untersuchung wird die Schätzperiode von 202 bis 2 Handelstage vor dem Earnings Call für die Schätzung der normalen Renditen herangezogen. Die Spanne von 200 Handelstagen entspricht hierbei der gängigen Praxis in der Ereignisstudienliteratur.¹⁴⁷

Im Rahmen der vorliegenden Arbeit werden logarithmierte Renditen berücksichtigt. Diese werden gemäß der Formel:

$$R_{i,t} = \ln \left(\frac{\text{Schlusskurs}_{i,t}}{\text{Schlusskurs}_{i,t-1}} \right) \quad (4)$$

berechnet, wobei $\text{Schlusskurs}_{i,t}$ den Schlusskurs der Aktie am Tag t und $\text{Schlusskurs}_{i,t-1}$ den einen Handelstag zuvor darstellt. Diese Methode bietet im Vergleich zu diskreten Renditen mehrere Vorteile: Logarithmierte Renditen sind mathematisch einfacher zu handhaben, zeigen bessere statistische Eigenschaften und neigen stärker zur Annäherung an eine Normalverteilung.¹⁴⁸ Aufgrund dieser Vorteile wird auch für Eventstudien häufig die Verwendung logarithmierter Renditen empfohlen.

Die kumulierte abnormale Rendite (CAR) bezeichnet die aufsummierte abnormale Rendite $AR_{i,t}$ eines Unternehmens i innerhalb eines definierten Ereignisfensters.¹⁴⁹ Die allgemeine Formel lautet:

$$CAR(t_1, t_2)_i = \sum_{t=t_1}^{t_2} AR_{i,t}, \quad (5)$$

wobei $t=0$ der Tag des Earnings Calls ist, t_1 die untere Grenze und t_2 die obere Grenze des Ereignisfensters bezeichnet. Für die vorliegende Untersuchung wird mit $CAR(-1,1)$ das initiale Ereignisfenster auf drei Tage um den Termin des Calls herum festgelegt. In Anlehnung an frühere Erkenntnisse, denen zufolge der Markt am Tag $t-1$ signifikant auf Unternehmensveröffentlichungen reagiert, beginnt das Ereignisfenster am Tag vor dem Call.¹⁵⁰ Die Verwendung eines dreitägigen Ereignisfensters bis einen Tag nach der Telefonkonferenz ist konsistent mit früheren Studien zur initialen Marktreaktion auf Informationen in Earnings Calls.¹⁵¹ Das Ereignisfenster umfasst anstelle von Kalendertagen Börsentage. Dies dient der Gewährleistung einer einheitlichen Definition des Zeitraums, innerhalb dessen der Kapitalmarkt in der Lage ist, neue Informationen zu verarbeiten.¹⁵²

Des Weiteren wird neben dem initialen Reaktionsfenster auch die nachfolgende Driftperiode in die Betrachtung einbezogen. Ball und Brown (1968) zeigten, dass die CAR nach Gewinnankündigungen bei als „gut“ klassifizierten Nachrichten im Zeitverlauf weiter ansteigen, während sie bei als „schlecht“ bewerteten Nachrichten weiter sinken.¹⁵³ Dieses Phänomen wird als Drift nach der Gewinnankündigung (Post-Earnings Announcement Drift, PEAD) bezeichnet. Die Resultate langfristiger Ereignisstudien legen die Vermutung nahe, dass die anfänglichen Marktreaktionen veröffentlichte Unternehmensinformationen nicht vollständig widerspiegeln.¹⁵⁴ Die Driftperiode auf den Kapitalmärkten ist der Zeitraum nach einer Ereignisankündigung, in dem Kursbewegungen weiter anhalten oder sich gegenläufig entwickeln können. Während die Märkte anfänglich schnell auf neue Informationen reagieren, verändern sich in der anschließenden Driftperiode die Auswirkungen des Ereignisses durch fortlaufende Informationsverarbeitung und Marktpsychologie. Diese Phase ist entscheidend, um die Effektivität der Informationsverarbeitung durch den Markt sowie die zeitliche Entwicklung der Einschätzungen der Investoren zu verstehen. Zu diesem Zweck werden in dieser Arbeit die kumulierten abnormalen Renditen $CAR(2,42)$ zwischen dem zweiten und 42. Handelstag nach den Earnings Calls analysiert.

3.4 Einflussfaktoren auf Aktienrenditen während Earnings Calls

Für vier Unternehmen in der Stichprobe konnten keine Kontrollvariablen ermittelt werden, sodass insgesamt 230 Unternehmen für die weitere Analyse verbleiben. Des Weiteren liegen teilweise für die verbleibenden Unternehmen keine Rendite- oder Kontrolldaten in Bezug auf einzelne Earnings Calls vor, sodass die weitere Analyse auf 4.190 Beobachtungen basiert.

Die grundlegende Struktur dieser Untersuchung basiert auf einer Regressionsanalyse, bei der CAR um die Zeit des Earnings Calls als abhängige Variable definiert werden. Die zentrale unabhängige Variable ist das mit dem Lexikon l ermittelte $Sentiment_{i,q,l}$ in dem Earnings Call von Unternehmen i zum Quartal q :

$$CAR(t_1, t_2)_{i,q} = \alpha + \beta * Sentiment_{i,q,l} + \sum Controls_{i,q} + \delta_i + \lambda_q + \varepsilon_{i,q}. \quad (6)$$

Zur Evaluierung der Reaktion von Investoren hinsichtlich des Tons eines Konferenzgesprächs erfolgt eine Schätzung von Panelregressionen mit festen Effekten für Unternehmen δ_i und Quartale λ_q . Des Weiteren wird der Ansatz der robusten Standardfehler nach White verfolgt, um potenzielle Heteroskedastizität in den Daten adäquat zu berücksichtigen.

¹⁴⁷Vgl. Holler (2016, pp. 38f).

¹⁴⁸Vgl. Gehrke (2022, p. 279); Holler (2016, p. 80).

¹⁴⁹Vgl. Gehrke (2022, pp. 283f).

¹⁵⁰Vgl. Henry (2008, p. 358).

¹⁵¹Vgl. Davis et al. (2014, p. 669); Price et al. (2012, p. 998).

¹⁵²Vgl. Gerpott (2009, p. 214).

¹⁵³Vgl. Ball und Brown (1968, p. 170).

¹⁵⁴Vgl. H. K. Baker und Nofsinger (2010, p. 338).

Auf die Berücksichtigung robuster Standardfehler für Autokorrelation wird verzichtet, da der Wooldridge-Test keinen Hinweis auf serielle Korrelation in den Residuen liefert.

Der Vektor $\sum Controls_{i,q}$ stellt eine Sammlung von Variablen auf Unternehmensebene zum Ende des jeweiligen Quartals q dar, auf das sich der Earnings Call bezieht. Durch die Aufnahme von Kontrollvariablen wird die Wahrscheinlichkeit verringert, dass die auf Wörterbüchern basierenden Sentiment-Maße Informationen erfassen, die mit anderen beobachtbaren Unternehmensmerkmalen am Datum der Telefonkonferenz zusammenhängen. Die Gewinnüberraschung bezeichnet die Differenz zwischen dem ausgewiesenen Gewinn eines Unternehmens und den Erwartungen des Marktes.¹⁵⁵ Sie stellt einen der maßgeblichen Einflussfaktoren bei der Bekanntgabe von Unternehmensergebnissen dar. Im Rahmen der vorliegenden Arbeit erfolgt die Ermittlung der Gewinnüberraschung mittels einer saisonalen Messung. Die Gewinnüberraschung eines Unternehmens i im Quartal q lässt sich demnach wie folgt berechnen:

$$SURP_{i,q} = \frac{(EPS_{i,q} - EPS_{i,q-4})}{Schlusskurs_{i,q-4}}, \quad (7)$$

wobei $EPS_{i,q}$ den Gewinn je Aktie des aktuellen Quartals, $EPS_{i,q-4}$ den Gewinn je Aktie des entsprechenden Vorjahresquartals und $Schlusskurs_{i,q-4}$ den Schlusskurs der Aktie zum Ende des betreffenden Quartals bezeichnet. Diese Vorgehensweise wird gegenüber einer auf Analystenprognosen basierenden Ermittlung bevorzugt, da der Gewinn im gleichen Quartal des Vorjahres aus der Perspektive des Managements eine bedeutsamere Benchmark darstellt als die durchschnittlichen Schätzungen der Analysten und letztere in der Regel Gewinne prognostizieren, die nicht im Einklang mit standardisierten Rechnungslegungsgrundsätzen stehen.¹⁵⁶ Die Verwendung einer saisonalen Messung der Gewinnüberraschung steht darüber hinaus im Einklang mit anderen Studien, die das Sentiment in Earnings Calls analysieren, und ermöglicht eine konsistente Vergleichbarkeit der Ergebnisse.¹⁵⁷

SIZE bezieht sich auf die Unternehmensgröße, die anhand der logarithmierten Marktkapitalisierung in Millionen US-Dollar gemessen wird. Mit steigender Unternehmensgröße nimmt die Menge an verfügbaren Informationen zu, wobei gleichzeitig eine potenziell stärkere regulatorische und öffentliche Kontrolle besteht.¹⁵⁸ Das BM-Verhältnis (BM) misst das Verhältnis des Buchwerts des Eigenkapitals zu der Marktkapitalisierung eines Unternehmens. Ein höheres Verhältnis deutet auf eine Unterbewertung des Unternehmens hin, während ein niedrigeres Verhältnis auf ein wachstumsorientiertes Unternehmen hinweist.¹⁵⁹ Mit

dem Return on Assets (ROA) lässt sich die Effektivität eines Unternehmens bei der Nutzung seiner Vermögenswerte zur Gewinnmaximierung beurteilen, was besonders bei nicht profitablen Unternehmen von Bedeutung ist. Ermittelt wird er durch Division des Nettogewinns durch die Gesamtsumme der Aktiva und anschließende Multiplikation des Ergebnisses mit 100. Die Kennzahl LEVERAGE gibt das Verhältnis von Gesamtschulden zu den jeweiligen Gesamtvermögenswerten eines Unternehmens an. Ein hoher Wert indiziert ein erhöhtes Insolvenzrisiko, während eine moderate Verschuldung auf eine effektive Nutzung des Fremdkapitals für Unternehmenswachstum hinweist.¹⁶⁰ LOSS ist eine Dummy-Variable, die den Wert 1 annimmt, sofern das Unternehmen im betrachteten Quartal einen Verlust erwirtschaftet hat, und den Wert 0 in allen anderen Fällen. Unternehmen, die Verluste verzeichnen, werden in der Regel als riskanter eingestuft, was sich in einer negativen Rendite niederschlagen kann.¹⁶¹ GROWTH bezeichnet das prozentuale Umsatzwachstum eines Unternehmens im Vergleich zum vorherigen Quartal. Die Variable wird zur Glättung von Extremwerten sowie zur Erzielung einer symmetrischeren Verteilung als natürlicher Logarithmus von eins plus Umsatzwachstumsrate berechnet. Ein signifikantes Umsatzwachstum kann als Indikator für eine positive Geschäftsentwicklung sowie vielversprechende Zukunftsaussichten gewertet werden.¹⁶² Die Jahre seit dem Börsengang werden durch FIRMAGE erfasst, berechnet als natürlicher Logarithmus von eins plus der Anzahl der Jahre seit dem Börsengang. In den ersten Jahren nach dem Börsengang eines Unternehmens sind historische Finanzdaten und Informationen für die Öffentlichkeit oft nicht verfügbar.¹⁶³ Dies resultiert in einer höheren Informationsasymmetrie sowie einem gesteigerten Risiko. Die Anzahl an Analysten, die an einer Konferenz teilnehmen, wird von ANALYSTS angegeben. Eine hohe Beteiligung von Analysten lässt auf ein großes Interesse am Unternehmen und dessen Bedeutung für die Finanzmärkte schließen.¹⁶⁴ Zudem lässt sich beobachten, dass die Teilnahme von Analysten den Informationsfluss und die Qualität, der in den Earnings Calls vermittelten Informationen beeinflusst.

4 Ergebnisse der Ereignisstudie zum Sentiment in Earnings Calls

4.1 Sentiment in Earnings Calls

4.1.1 Deskriptive Statistiken des Sentiments in Earnings Calls

Die Gegenüberstellung der Median- und Mittelwerte der drei Sentimentvariablen zeigt Unterschiede in der Stimmungsermittlung zwischen den drei verschiedenen Lexika. Das LM-Lexikon weist mit einem Medianwert von

¹⁵⁵Vgl. Price et al. (2012, p. 998).

¹⁵⁶Vgl. Price et al. (2012, p. 998).

¹⁵⁷Vgl. Milian et al. (2017, p. 70); Blau et al. (2015, p. 208); Brockman et al. (2015, p. 26); Price et al. (2012, p. 998).

¹⁵⁸Vgl. Bochkay et al. (2019, p. 49).

¹⁵⁹Vgl. Price et al. (2012, p. 999).

¹⁶⁰Vgl. Price et al. (2012, p. 999).

¹⁶¹Vgl. Mather et al. (2021); Davis et al. (2014, p. 657).

¹⁶²Vgl. Yamamoto et al. (2022, p. 249); Bochkay et al. (2019, p. 53); Blau et al. (2015, p. 214); Davis et al. (2014, p. 657).

¹⁶³Vgl. Ning et al. (2023).

¹⁶⁴Vgl. Cicon (2015, p. 72).

-0,0137 auf ein insgesamt leicht negatives Sentiment in den Earnings Calls hin, was auf eine tendenziell pessimistische Grundstimmung schließen lässt. Im Gegensatz dazu zeigen das Harvard-GI-Lexikon und das Henry-Lexikon mit Medianwerten von 0,09561 bzw. 0,5875 eine überwiegend

positive Stimmung. Die durchschnittlichen monatlichen Sentiment-Werte je Lexikon im Betrachtungszeitraum sind in Abbildung 1 dargestellt und veranschaulichen die Unterschiede zwischen den Lexika.

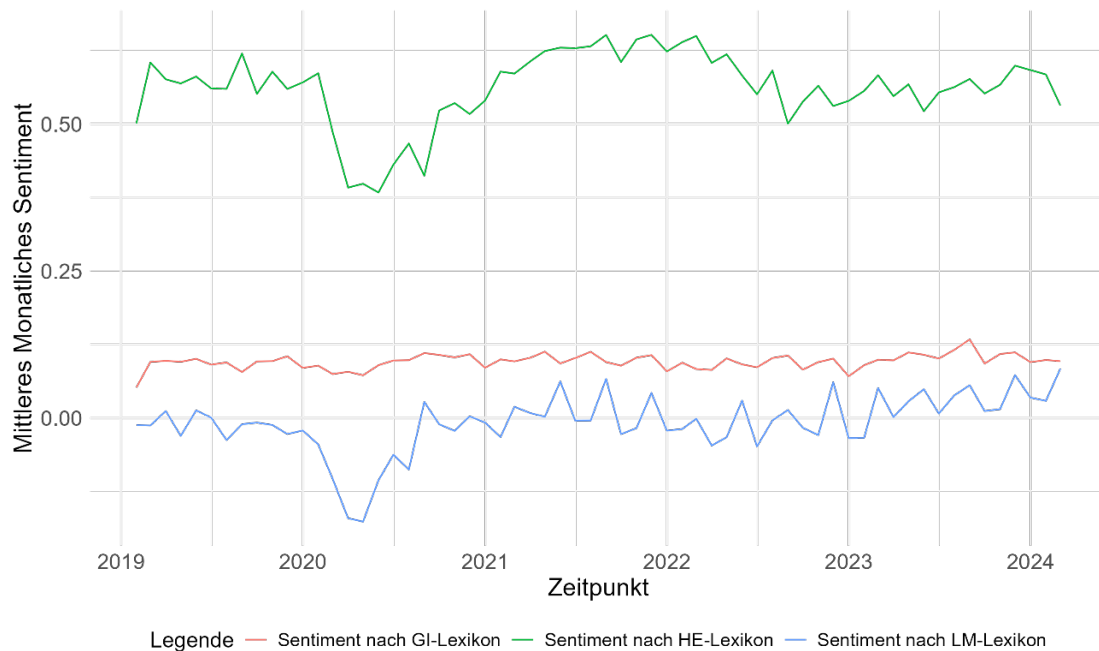


Abbildung 1: Mittleres Monatliches Sentiment in Earnings Calls nach Loughran McDonald (LM), Henry (HE) und Harvard-GI (GI) Lexikon

Die Betrachtung der Spannweite der Sentiment-Werte verdeutlicht, dass die Stimmung in den Earnings Calls sowohl starke negative als auch starke positive Ausprägungen annimmt. Der niedrigste gemessene Wert aller Lexika von -0,6238 wurde mit dem LM-Lexikon für einen Earnings Call des ersten Quartals 2022 von Medpace Holdings ermittelt. Obwohl die Gewinnerwartungen leicht übertroffen wurden, fokussierte sich die Diskussion in der Konferenz primär auf Verzögerungen und Stornierungen von Programmen, was Besorgnis hinsichtlich der künftigen Wachstumspläne auslöste. Im Gegensatz dazu wurde der höchste Sentiment-Wert von 0,9080 für einen Earnings Call des ersten Quartals 2021 von West Pharmaceutical Services Inc. mit dem Henry-Lexikon erfasst. Im Rahmen der Veranstaltung wurden Informationen zu einer hohen Nachfrage nach den Produkten in Verbindung mit der Coronavirus-Pandemie diskutiert.

Die Unterschiede in den deskriptiven Statistiken lassen sich in großen Teilen auf die unterschiedlichen Verhältnisse von positiven zu negativen Begriffen in den Lexika zurückführen. Im LM-Lexikon liegt dieses Verhältnis bei 1:6,65, was bedeutet, dass es wesentlich mehr negative als positive Wörter umfasst. Dies führt zu den niedrigsten und negativsten gemessenen Sentiment-Werten unter den untersuchten Lexika. Demgegenüber zeigen das GI- und das HE-Lexikon ein aus-

gewogeneres bzw. umgekehrtes Verhältnis von positiven zu negativen Begriffen. Im GI-Lexikon ist die Anzahl der negativen Begriffe leicht höher als die Anzahl der positiven Begriffe (Verhältnis 1:1,22). Im Gegensatz dazu zeigt sich im HE-Lexikon ein leichtes Überwiegen positiver Begriffe (Verhältnis 1,24:1). Diese unterschiedlichen Verhältnisse manifestieren sich in einer positiveren Tendenz hinsichtlich der gemessenen Sentiment-Werte. Darüber hinaus werden die geringsten Standardabweichungen im GI-Lexikon beobachtet, was auf eine konsistentere Bewertung des Sentiments hinweist. Demgegenüber weisen das LM- und das Henry-Lexikon höhere Abweichungen auf, was auf eine differenziertere Analyse zurückzuführen sein könnte. Die Varianz der Sentiment-Werte in Abhängigkeit vom verwendeten Lexikon verdeutlicht die Bedeutung der spezifischen Eigenschaften sowie des Gewichtungsverhältnisses der Lexika bei der Analyse der Beziehung zwischen Sentiment und Marktreaktionen. Die deskriptiven Statistiken legen insgesamt nahe, dass die Auswahl des Lexikons einen signifikanten Einfluss auf die gemessenen Sentiment-Werte hat, was bei der weiteren Analyse der abnormalen Renditen zu berücksichtigen ist.

In Anbetracht der unterschiedlichen Inhalte und Sprecherperspektiven ist es entscheidend, das jeweilige Sentiment des Earnings Calls differenziert zu betrachten. Hierbei ist insbesondere eine Unterscheidung zwischen der

allgemeinen Stimmung während des Calls und den spezifischen Segmenten wie Präsentation und Q&A oder Rollen wie Manager und Analysten erforderlich. Im Henry-Lexikon wird die Präsentation tendenziell positiver bewertet als das Q&A (Median: 0,5966 vs. 0,5789). In diesem Lexikon fällt der Unterschied zwischen den beiden Abschnitten jedoch weniger deutlich aus als in den anderen untersuchten Lexika. Der im Vergleich zur Präsentation negativere Q&A-Abschnitt ist konsistent mit ähnlichen Befunden in der Literatur, wonach diese Beobachtung nicht auf das Henry-Lexikon beschränkt ist.¹⁶⁵ Das Beantworten kritischer und detaillierter Fragen während der Q&A-Sitzung kann demzufolge dazu führen, dass das Management zu sponta-

nen und ehrlichen Antworten gezwungen wird, die weniger optimistisch ausfallen als ihre vorbereiteten Präsentationen. Entgegen dieser Erwartung weisen das Harvard-GI- und das LM-Lexikon in dieser Arbeit eine positivere Tendenz im Frage-Antwort-Abschnitt (Median LM: 0,0698, GI: 0,1039) als in der Präsentation (Median LM: -0,0988, GI: 0,0887) auf. Dies deutet darauf hin, dass die Fragen der Analysten selbst eine positivere Formulierung aufweisen oder dass die Unternehmensführung während der Q&A-Sitzung optimistischere oder beruhigendere Antworten als in der Präsentation gibt.

Tabelle 3: Deskriptive Statistiken

	Variabel	Mittelwert	Median	Min.	Max.	SD
Sentiment	Call (LM)	-0.0190	-0.0137	-0.6238	0.5190	0.1511
	Call (HE)	0.5644	0.5875	-0.1341	0.9080	0.1396
	Call (GI)	0.0967	0.0956	-0.0821	0.3461	0.0558
	Präsentation (LM)	-0.1081	-0.0988	-0.8272	0.4648	0.1822
	Präsentation (HE)	0.5654	0.5966	-0.1833	0.9192	0.1751
	Präsentation (GI)	0.0900	0.0887	-0.1463	0.3702	0.0666
	Q&A (LM)	0.0675	0.0698	-0.6403	0.6059	0.1748
	Q&A (HE)	0.5642	0.5789	-0.1023	1.0000	0.1465
	Q&A (GI)	0.1044	0.1039	-0.1137	0.3787	0.0676
	Manager Q&A (LM)	0.1024	0.1060	-0.6552	0.6897	0.1850
	Manager Q&A (HE)	0.6171	0.6309	-0.0909	1.0000	0.1513
	Manager Q&A (GI)	0.1043	0.1053	-0.1644	0.4118	0.0787
	Analyst Q&A (LM)	-0.0108	0.0000	-1.0000	1.0000	0.2691
	Analyst Q&A (HE)	0.4094	0.4254	-0.6667	1.0000	0.2237
	Analyst Q&A (GI)	0.0943	0.0934	-0.3247	0.5023	0.0881
Kontrollvariablen	SURP	0.0044	0.0010	-0.4215	1.4319	0.0512
	SIZE	8.7316	8.5499	4.8819	14.8433	1.5387
	BM	0.4394	0.3318	-2.7550	8.7767	0.5640
	ROA	0.0518	0.0504	-0.6283	0.5078	0.0775
	LEVERAGE	0.2863	0.2764	0.0000	2.2923	0.1963
	GROWTH	0.0164	0.0183	-4.1044	3.6384	0.2354
	LOSS	0.1649	0.0000	0.0000	1.0000	0.3711
	FIRMAGE	3.1094	3.2594	0.7396	4.6509	0.7854
CAR	ANALYSTS	7.6737	7.0000	1.0000	24.0000	3.7025
	CAR(-1,1)	0.0010	0.0042	-0.4587	0.5630	0.0869
	CAR(2,42)	-0.0060	-0.0081	-1.2156	1.2957	0.1446

Innerhalb des Frage-und-Antwort-Abschnitts weisen alle drei Lexika einen positiveren Trend in den Sentiment-Werten der Aussagen von Managern in dem Q&A im Vergleich zu den Aussagen von Analysten auf. Die Medianwerte des Manager-Sentiments sind höher als die der Analysten-Sentiments. Insbesondere im Henry-Lexikon zeigt sich eine ausgeprägte Differenz, während der Unterschied im GI-Lexikon weniger deutlich ist. Diese Beobachtung differierender Haltungen zwischen Managern und Analysten deckt

sich mit den Resultaten vorheriger Studien.¹⁶⁶ Die hohe Streuung der Sentiment-Werte der Analysten im Q&A-Teil deutet darauf hin, dass Analysten während des Q&A-Teils stark unterschiedliche Einschätzungen und Reaktionen aufweisen. Dabei legt die positivere Tonalität der Manager im Q&A-Abschnitt im Vergleich zu den Analysten nahe, dass Unternehmensvertreter versuchen, negative Fragestellungen durch eine optimistische Argumentation zu kompensieren. Des Weiteren lässt sich eine positivere Stimmung in den

¹⁶⁵Vgl. De Amicis et al. (2021); Fu et al. (2019, p. 649); Borochin et al. (2018, p. 81); Brockman et al. (2015, p. 27)

¹⁶⁶Vgl. Fu et al. (2019, p. 649); Borochin et al. (2018, p. 81); Brockman et al. (2015, p. 27)

Aussagen der Manager in den Q&A-Abschnitten aller drei Lexika im Vergleich zu ihren Aussagen in der Präsentation feststellen. Dies deutet darauf hin, dass Manager während tendenziell offener und positiver auf Fragen eingehen als in der formalen Präsentation, was zu einem insgesamt höheren Sentiment-Wert in diesem Abschnitt des Calls führt.

Die deskriptiven Statistiken für alle abhängigen, erklärenden und Kontrollvariablen sind in Tabelle 3 dargestellt.

Die anfänglichen Marktreaktionen CAR(-1,1) um den Tag des Earnings Calls variieren zwischen -0,4587 und 0,5630. Diese Spannweite lässt auf Unterschiede in der Relevanz der veröffentlichten Informationen sowie divergierende Erwartungen der Investoren schließen. Der ermittelte Median von 0,0042 deutet auf tendenziell positive Marktreaktionen hin. Im Gegensatz dazu lässt die Driftperiode CAR(2,42) nach den Earnings Calls mit einem Median von -0,0081 eine leicht negative Tendenz erkennen. Die signifikante Spannweite demonstriert, dass einige Unternehmen in dieser Periode erhebliche Verluste oder Gewinne verzeichnen. Ein Vergleich der Werte von CAR(1,1) mit denen von CAR(2,42) zeigt, dass die letztgenannten Werte etwa doppelt so hoch sind. Dies lässt den Schluss zu, dass der Informationsgehalt der Earnings Calls während der anfänglichen Reaktionszeit nicht vollständig in den Kursen enthalten ist. Die tendenziell positiven direkten Reaktionen auf Earnings Calls und die oft negative nachfolgende Performance weisen auf eine komplexe Dynamik zwischen kurzfristigen Marktreaktionen und längerfristigen Entwicklungen hin.

Die Analyse der Kontrollvariablen zeigt, dass der Median der Gewinnüberraschungen bei 0,0010 liegt, was auf eine leichte Tendenz zu positiven Überraschungen hindeutet, die mit früheren Studien übereinstimmt.¹⁶⁷ Das BM-Verhältnis beträgt 0,4394, der Return on Assets 0,0518 und der Verschuldungsgrad 0,2863. Die durchschnittliche logarithmierte Marktkapitalisierung der untersuchten Unternehmen beträgt 8,7316 Millionen US-Dollar. Das Umsatzwachstum variiert zwischen -4,1043 und 3,6384, was eine breite Spanne an Wachstumsraten impliziert. Das durchschnittliche Alter der betrachteten Unternehmen, berechnet als natürlicher Logarithmus von 1 plus der Jahre seit dem Börsengang, beträgt 3,1094. Die Mehrheit der Unternehmen ist folglich bereits seit mehreren Jahren börsennotiert, während einige ihren Börsengang während des Beobachtungszeitraums durchgeführt haben. Die Anzahl der an den Calls teilnehmenden Analysten variiert zwischen einem und 24, wobei ein Durchschnitt von etwa acht Analysten pro Call zu verzeichnen ist.

4.1.2 Korrelationen des Sentiments in Earnings Calls

Wie der Korrelationsmatrix in Anhang 1 zu entnehmen sind die Korrelationen zwischen den Tonmaßen der verschiedenen Lexika durchweg positiv, jedoch nicht perfekt. Dies lässt den Schluss zu, dass die Lexika zum einen teilweise ähnliche, zum anderen aber auch unterschiedliche

Informationen zur Stimmung erfassen. Besonders hohe Übereinstimmungen zeigen sich zwischen dem LM- und dem Henry-Lexikon, insbesondere bei der Erfassung des Gesamt-Sentiments im Call und während der Präsentation. Des Weiteren ist eine Varianz der Lexika in Bezug auf das Korrelationsniveau zwischen den ermittelten Tonmaßen und den kumulierten abnormalen Renditen im initialen Reaktionsfenster sowie der Driftperiode festzustellen. Die Korrelationen der Lexika mit CAR(-1,1) sind durchweg positiv. Das Henry-Lexikon weist hierbei die höchsten Werte auf. Eine Ausnahme bildet die höchste Gesamtkorrelation zwischen dem Sentiment der Analysten im Q&A, ermittelt mit dem LM-Lexikon, und CAR(-1,1). In der Driftperiode sind die Korrelationen des Sentiments mit CAR(2,42) insgesamt geringer und teilweise negativ.

Die Analyse der entsprechenden Tonmaße ergibt, dass in den verschiedenen Abschnitten und bei den unterschiedlichen Sprechern innerhalb der Earnings Calls sprachliche Nuancen zu beobachten sind. Die positiven Korrelationen zwischen den Tonvariablen in verschiedenen Abschnitten (Präsentation, Q&A) und Rollen der Redner (Unternehmensvertreter, Analysten) bei allen drei Lexika lassen den Schluss zu, dass eine gemeinsame Signalkomponente existiert, die die übergeordnete Stimmung während der gesamten Telefonkonferenz reflektiert. Diese gemeinsame Komponente ist jedoch nicht perfekt, da die Korrelationen zwischen den verschiedenen Abschnitten und Rednerrollen mit gleichem Lexikon zwar positiv sind, aber zwischen 0,2219 und 0,5107 variieren. Die geringsten Korrelationen sind zwischen dem Sentiment in der Präsentation und dem Sentiment der Analysten zu beobachten. Dies lässt den Schluss zu, dass Analysten und Management unterschiedliche Perspektiven und divergierende Erwartungen haben.

Das Sentiment im Q&A weist, mit Ausnahme der Analyse mit dem GI-Lexikon, im initialen Reaktionsfenster eine stärkere positive Korrelation mit CAR(-1,1) auf als das Sentiment in der Präsentation. Im Vergleich dazu lässt sich in der Driftperiode CAR(2,42) eine leicht stärkere negative Korrelation des Sentiments der Präsentation mit den abnormalen Renditen beobachten, was auf eine unterschiedlich starke Wirkung der beiden Abschnitte des Earnings Calls hindeutet. Eine Gegenüberstellung des Sentiments innerhalb des Q&A zeigt, dass das Sentiment der Analysten im initialen Reaktionsfenster eine stärkere Korrelation mit den CAR(-1,1) aufweist als das Sentiment der Manager. Das Analysten-Sentiment erzielt sowohl für das LM- als auch für das Henry-Lexikon die höchsten Korrelationswerte zu CAR unter allen untersuchten Variablen. In der Driftperiode CAR(2,42) lassen sich keine signifikanten Unterschiede zwischen den Einschätzungen der Analysten und der Manager im Q&A feststellen.

Die Korrelationen zwischen den Kontrollvariablen und den abnormalen Renditen sind insgesamt von geringer Intensität, was darauf hindeutet, dass diese Variablen nur eine untergeordnete Rolle bei der Erklärung der abnormalen Renditen spielen. Darüber hinaus lässt die geringe Korrelation zwischen Sentiment und Kontrollvariablen den Schluss zu, dass das Sentiment unabhängige Informationen liefert.

¹⁶⁷Vgl. Milian et al. (2017, p. 70); Blau et al. (2015, p. 209); Brockman et al. (2015, p. 27); Price et al. (2012, p. 1000).

In Anbetracht der Größenordnungen der meisten Korrelationen ist davon auszugehen, dass Multikollinearität bei der Auswertung der nachfolgenden Regressionen keine wesentlichen Schwierigkeiten bereiten wird. Eine Betrachtung der Korrelationsmatrix zeigt, dass keine starken Korrelationen zwischen den unabhängigen Variablen beobachtet werden. Lediglich zwischen Variablen, die nicht gleichzeitig in den Regressionen berücksichtigt werden, treten starke

Korrelationen von über 0,7 auf. Dies betrifft beispielsweise das Sentiment des gesamten Calls und das der Präsentation. Die stärksten Korrelationen zwischen gleichzeitig betrachteten Variablen bestehen zwischen den Kontrollvariablen ROA und LOSS mit einem Wert von -0,53 sowie zwischen der Unternehmensgröße und der Anzahl der teilnehmenden Analysten.

Tabelle 4: Einfluss des Gesamtsentiments in Earnings Calls sowie des differenzierten Sentiments in Präsentation und Q&A auf CAR(-1,1)

	LM	HE	GI	LM	HE	GI
Gesamter Call	0.136*** (0.0144)	0.148*** (0.0157)	0.281*** (0.0439)			
Präsentation				0.054*** (0.0141)	0.066*** (0.0134)	0.153*** (0.0341)
Q&A				0.067*** (0.0128)	0.073*** (0.0141)	0.103*** (0.0309)
R ²	0.0737	0.0809	0.0630	0.0812	0.0888	0.0722
Adjusted R ²	0.0116	0.0193	0.0002	0.0132	0.0214	0.0035

*** $p < 0.1\%$ ** $p < 1\%$ * $p < 5\%$. $p < 10\%$

4.2 Einfluss des Sentiments in Earnings Calls auf kumulierte abnormale Renditen

4.2.1 Gesamtsentiment

Im Rahmen dieser Untersuchung wurden Panelregressionen durchgeführt, um den Zusammenhang zwischen kumulierten abnormalen Renditen im initialen Ereignisfenster CAR(-1,1) und dem Sentiment in Earnings Calls zu analysieren. Das Sentiment wurde jeweils anhand der Lexika LM, Henry und Harvard-GI in einem ersten Schritt für den gesamten Earnings Call ermittelt:

$$CAR(-1,1)_{i,q} = \alpha + \beta * Sentiment_{Gesamt,i,q,l} + \sum Controls_{i,q} + \delta_i + \lambda_q + \varepsilon_{i,q}. \quad (8)$$

Die in Tabelle 4 dargestellten Ergebnisse aller Modelle legen nahe, dass ein positiveres Sentiment einen signifikant positiven Einfluss (auf 0,1%-Niveau) auf die CAR(-1,1) ausübt. Auf dieser Basis lässt sich die Hypothese H1a bestätigen, dass ein positives Sentiment in Earnings Calls signifikant positive Marktreaktionen hervorruft, während ein negatives Sentiment zu signifikant negativen Marktreaktionen führt. Dabei verdeutlicht die Konsistenz der Ergebnisse, dass das Sentiment in Earnings Calls verlässlich und robust die abnormalen Renditen im initialen Reaktionsfenster beeinflusst. Dies deutet darauf hin, dass das Sentiment von den Investoren systematisch verarbeitet wird und nicht nur zufällig mit den Renditen korreliert. Das Bestimmtheitsmaß fällt mit Werten zwischen 0,0630 und 0,0809 gering aus, was darauf hinweist, dass lediglich ein geringer Teil der Varianz der abnormalen Renditen durch die unabhängigen Variablen erklärt wird. Diese Tatsache ist bei Finanzmarktdaten, die in

der Regel von einer Vielzahl an schwer quantifizierbaren Faktoren beeinflusst werden, nicht ungewöhnlich und konsistent zu anderen Arbeiten zu Earnings Calls.¹⁶⁸ Die höchste Erklärungskraft bezüglich des Sentiments weist das mit dem Henry-Lexikon ermittelte Modell auf, während das Modell des Harvard-GI Lexikons das niedrigste Bestimmtheitsmaß zeigt. Des Weiteren verzeichnet das Modell des Harvard-GI Lexikons eine höhere Standardabweichung für den Koeffizienten, als die beiden anderen Lexika, was auf eine größere Unsicherheit der Schätzung hinweist.

Der Koeffizient des Harvard-GI-Lexikons ist mit 0,281 der höchste unter den untersuchten Lexika und deutet zunächst auf einen starken positiven Effekt hin. Der hohe Wert könnte jedoch auf die geringere Spannweite der Sentiment-Werte in diesem Lexikon zurückzuführen sein, wodurch geringe Änderungen eine größere relative Auswirkung haben. Ein hoher Koeffizient indiziert eine hohe Sensitivität des Lexikons, wobei jedoch zu berücksichtigen ist, dass dies auch auf weniger differenzierte Sentiment-Werte zurückzuführen sein könnte. Diese Feststellung schmälert nicht zwangsläufig die Eignung des Lexikons, sondern demonstriert, dass der Effekt des Sentiments pro Einheit potenziell größer sein kann.

Die Sentimentvariablen erweisen sich nicht nur als signifikant, sondern zeigen zudem höhere Koeffizienten als andere bedeutsame Einflussfaktoren, wie beispielsweise die Unternehmensgröße, der Verschuldungsgrad oder das Umsatzwachstum. Andere Kontrollvariablen, wie beispielsweise der Return on Assets (ROA) sowie die Anzahl der an dem Call teilnehmenden Analysten, weisen demgegenüber

¹⁶⁸Vgl. De Amicis et al. (2021); Brockman et al. (2015, p. 36); Cicon (2015, p. 75); Huang et al. (2013, p. 1006); Price et al. (2012, p. 1005).

keinen signifikanten Einfluss auf. Die Effekte des Sentiments sind insgesamt konsistenter, signifikanter und stärker als die der Kontrollvariablen. Dies demonstriert, dass die qualitativ und emotional geprägten Sentimentfaktoren in der initialen Marktreaktion einen maßgeblicheren Einfluss auf die abnormalen Renditen ausüben als rein quantitative Finanzkennzahlen.

4.2.2 Separater Einfluss von Präsentations- und Q&A-Sentiment

In einem nächsten Schritt wurde anstelle des Sentiments des gesamten Calls, das Sentiment der Präsentation $Sentiment_{Präsentation,i,q,l}$ und des Q&As $Sentiment_{Q\&A,i,q,l}$ einzeln betrachtet:

$$\begin{aligned} CAR(-1, 1)_{i,q} = & \alpha + \beta_1 * Sentiment_{Präsentation,i,q,l} \\ & + \beta_2 * Sentiment_{Q\&A,i,q,l} \\ & + \sum Controls_{i,q} + \delta_i + \lambda_q + \varepsilon_{i,q}. \end{aligned} \quad (9)$$

Die Koeffizienten für das Sentiment der Präsentation sowie des Q&A sind durchgehend positiv und auf dem 0,1-%-Niveau signifikant (siehe Tabelle 4). Die Ergebnisse zeigen, dass sowohl das Sentiment der Präsentation als auch das Sentiment des Q&A-Abschnitts signifikante und unabhängige Einflüsse auf die abnormalen Renditen ausüben. Damit wird die Hypothese H2a gestützt. Für das LM- und Henry-Lexikon weist das Q&A leicht höhere Koeffizienten als die Präsentation auf, was auf einen stärkeren Einfluss des Frage- und Antwort-Abschnitts auf die CAR(-1,1) hindeutet und die Hypothese 2b bestätigt. Die höheren Koeffizienten im Q&A-Abschnitt legen nahe, dass Investoren auf die spontanen Antworten des Managements oder Fragen der Analysten stärker reagieren als auf die vorbereiteten Präsentationen. Dies könnte darauf zurückzuführen sein, dass sie diese Informationen als authentischer und aufschlussreicher für die zukünftige Unternehmensentwicklung einschätzen. Im Gegensatz dazu weist das GI-Lexikon höhere Koeffizienten für den Präsentationsteil auf, was eine gegenteilige Beziehung impliziert.

Eine separate Betrachtung des Sentiments der Präsentation und des Q&As führt zu einer leichten Verbesserung des adjustierten Bestimmtheitsmaßes der Modelle mit LM- und Henry-Lexikon. Diese detaillierte Betrachtung liefert folglich eine bessere Erklärung der abnormalen Renditen als das aggregierte Modell, obwohl sie eine größere Anzahl an Prädiktoren beinhaltet. Die signifikanten und differenzierten Einflüsse der einzelnen Abschnitte des Earnings Calls könnten in dem aggregierten Modell, das lediglich das Sentiment des gesamten Calls zusammenfasst, verwischt werden. Dies verdeutlicht die Relevanz einer differenzierten Betrachtung der verschiedenen Segmente von Earnings Calls. Wie bereits in der vorherigen Analyse zeigt auch diesmal das Henry-Lexikon die beste Modellgüte, während das GI-Lexikon gegenläufige Ergebnisse liefert. Des Weiteren lässt sich auch in diesem Modell anhand des Standardfehlers

des Harvard-GI-Lexikons ableiten, dass die Schätzungen weniger präzise ausfallen als bei LM und Henry.

4.2.3 Differenzierter Einfluss von Manager- und Analysten-Sentiment im Q&A

Im Anschluss wurde neben der Präsentation anstelle des Sentiments des gesamten Q&As mit $Sentiment_{AnalystQ\&A,i,q,l}$ das Sentiment der Analysten und $Sentiment_{ManagerQ\&A,i,q,l}$ das der Manager in der Frage- und Antwortsektion einzeln betrachtet:

$$\begin{aligned} CAR(-1, 1)_{i,q} = & \alpha + \beta_1 * Sentiment_{Präsentation,i,q,l} \\ & + \beta_2 * Sentiment_{ManagerQ\&A,i,q,l} \\ & + \beta_3 * Sentiment_{AnalystQ\&A,i,q,l} \\ & + \sum Controls_{i,q} + \delta_i + \lambda_q + \varepsilon_{i,q}. \end{aligned} \quad (10)$$

In Bezug auf die Koeffizienten des Präsentations-Sentiments lässt sich, wie aus Tabelle 5 ersichtlich, mit allen drei Lexika ein positives und hochsignifikantes Ergebnis auf dem 0,1-%-Niveau verzeichnen. Die Koeffizienten der Präsentation übersteigen die der einzelnen Elemente des Q&A-Abschnitts. Der vorherige Vergleich hat gezeigt, dass das Q&A insgesamt eine größere Bedeutung hat. Bei einer differenzierteren Betrachtung wird jedoch deutlich, dass das Sentiment in den kürzeren Segmenten der Managerantworten und Analystenfragen weniger gewichtet wird als die Präsentation.

Eine separate Betrachtung des Sentiments innerhalb des Frage- und Antwort-Abschnitts verdeutlicht, dass das Analysten-Sentiment bei den LM- und Henry-Lexika ein hochsignifikantes Ergebnis auf dem 0,1-%-Niveau zeigt, wobei die Koeffizienten doppelt so hoch sind wie bei den Antworten des Managements, was die Hypothese H2d bestätigt. Diese Ergebnisse legen nahe, dass die höhere Bedeutung des Q&A-Abschnitts im Vergleich zur Präsentation insbesondere durch die Fragen der Analysten bedingt ist. Demgegenüber lässt sich mit dem Harvard-GI-Lexikon kein signifikanter Einfluss des Analyst-Q&A-Sentiments feststellen. Die Auswertung des Manager-Q&A-Sentiments ergibt signifikante Ergebnisse für das Harvard-GI-Lexikon (1-%-Niveau) und das LM-Lexikon (10-%-Niveau), während beim Henry-Lexikon keine Signifikanz nachgewiesen werden kann. Die vorliegenden Ergebnisse stützen die Hypothese H2c nur teilweise, da sowohl das Analysten-Sentiment als auch das Manager-Sentiment nicht durchgängig signifikante Einflüsse aufweisen.

Insgesamt deuten die Ergebnisse darauf hin, dass die Lexika verschieden auf die spezifischen und differenzierten Einflüsse der einzelnen Teile der Earnings Calls reagieren und unterschiedliche Fähigkeiten zur präzisen Erfassung der jeweiligen Segmente besitzen. Das Henry-Lexikon verzeichnet erneut die beste Modellgüte, während das Harvard GI-Lexikon ein negatives adjustiertes Bestimmtheitsmaß aufweist und somit keine höhere Erklärungskraft bietet als der Mittelwert der abhängigen Variable. Die Modellgüte für

Tabelle 5: Einfluss des Sentiments in der Präsentation sowie der Aussagen von Managern und Analysten im Q&A von Earnings Calls auf CAR(-1,1)

	LM	HE	GI	Komb.
Präsentation	0.056*** (0.0141)	0.070*** (0.0133)	0.160*** (0.0340)	0.081*** (0.0124)
Manager Q&A	0.018 (0.0106)	0.018 (0.0127)	0.079** (0.0247)	0.062* (0.0247)
Analyst Q&A	0.047*** (0.0065)	0.044*** (0.0082)	0.015 (0.0191)	0.049*** (0.0066)
R ²	0.0866	0.0885	0.0688	0.0966
Adj. R ²	0.0187	0.0208	-0.0003	0.0295

*** $p < 0.1\%$ ** $p < 1\%$ * $p < 5\%$. $p < 10\%$

die Präsentation und das Gesamtsentiment des Q&As ist für alle Lexika höher als bei einer separaten Betrachtung von Manager- und Analysten-Sentiment innerhalb des Frage- und Antwortabschnitts. Die fehlende Signifikanz einzelner Teile des Q&As deutet darauf hin, dass die unzureichende Eignung der Lexika für bestimmte Segmente der Earnings Calls dazu führt, dass nicht signifikante Variablen einbezogen werden, die die erklärte Varianz verringern.

4.2.4 Langfristige Auswirkungen des Sentiments

Im Zeitverlauf beginnen Investoren, die veröffentlichten Informationen gründlicher zu analysieren und die langfristigen Implikationen zu bewerten, was zu einer Anpassung der Marktreaktionen führen kann. Daher wird auch der Einfluss des Sentiments auf die kumulierten abnormalen Renditen CAR(2,42) zwei bis 42 Tage nach dem Earnings Call betrachtet.

In der Driftperiode zeigt das Sentiment aller Lexika einen signifikanten negativen Einfluss auf die abnormalen

Renditen, was darauf hindeutet, dass die anfängliche positive Marktreaktion nicht vollständig nachhaltig ist und eine negative Korrektur der anfänglichen Renditen eintritt (siehe Tabelle 6). Die Hypothese H1c wird somit bestätigt, wobei jedoch keine vollständige Umkehr eintritt. Diese Beobachtungen stehen im Einklang mit bestehenden Studien, die ähnliche Korrekturen belegen. So dokumentieren beispielsweise Huang et al. (2020) eine langfristige Revision der anfänglichen positiven Reaktion auf den Ton von Gewinnmitteilungen.¹⁶⁹ Druz et al. (2020) und Price et al. (2012) stellten hingegen fest, dass das Sentiment im Zeitverlauf an Bedeutung gewinnt und einen gleichgerichteten Drift wie die initiale Reaktion auslöst, was sie darauf zurückführen, dass qualitative Informationen schwerer zu verarbeiten sind und daher verzögert in die Marktbewertungen einfließen.¹⁷⁰

Tabelle 6: Einfluss des Sentiments in der Präsentation sowie der Aussagen von Managern und Analysten im Q&A von Earnings Calls auf CAR(2,42)

	LM	HE	GI	LM	HE	GI
Gesamter Call	-0.0617* (0.0249)	-0.0958*** (0.0246)	-0.1744*** (0.0598)			
Präsentation				-0.045* (0.0207)	-0.081*** (0.0209)	-0.173*** (0.0523)
Q&A				-0.025 (0.0183)	0.001 (0.0197)	-0.024 (0.0430)
R ²	0.1096	0.1125	0.1096	0.1147	0.1177	0.1154
Adjusted R ²	0.0499	0.0530	0.0500	0.0492	0.0524	0.0499

*** $p < 0.1\%$ ** $p < 1\%$ * $p < 5\%$. $p < 10\%$

Die Modellgüte ist in der Driftperiode höher als in der initialen Reaktion, wobei dieser Anstieg primär durch Kontrollvariablen bestimmt wird und das Sentiment eine untergeord-

nete Rolle spielt. Während die Gewinnüberraschung in der initialen Reaktion keine signifikante Auswirkung auf die abnormalen Renditen hat, lässt sich in der anschließenden Driftperiode ein signifikanter, negativer Effekt beobachten. Zudem sind in diesem Zeitraum signifikante Einflüsse der Unternehmensgröße sowie des Return on Assets festzustellen. Die Er-

¹⁶⁹Vgl. Huang et al. (2013, p. 1111).

¹⁷⁰Vgl. Druz et al. (2020, pp. 59, 66); Price et al. (2012, p. 994).

gebnisse deuten darauf hin, dass nach der initialen, emotional geprägten Reaktion eine verstärkte Fokussierung auf quantitative Unternehmenskennzahlen und langfristige Bewertungen stattfindet, bei der Investoren potenzielle Risiken und Unsicherheiten intensiver berücksichtigen und ihre ursprüngliche Reaktion teilweise revidieren. Dies unterstreicht die Notwendigkeit, sowohl qualitative als auch quantitative Aspekte in die Analyse von Unternehmenskommunikationen einzubeziehen, um ein umfassendes Verständnis der Marktreaktionen zu gewährleisten.

Während das Sentiment im Rahmen des Q&A (inklusive der darin enthaltenen Aussagen von Managern und Analysten) initial eine wesentliche Rolle spielt und $CAR(-1,1)$ beeinflusst, sind seine Auswirkungen in der späteren Driftperiode nicht signifikant. Die Präsentation hingegen zeigt einen signifikanten negativen Einfluss auf $CAR(2,42)$ in der Driftperiode nach dem Call. Diesbezüglich lassen sich zwei Erklärungsansätze anführen: Erstens könnte dies darauf hinweisen, dass das Sentiment der Präsentation in der Driftperiode aufgrund tiefergehender Bewertungen und der Erkennung langfristiger Risiken an Bedeutung gewinnt. Die Präsentation liefert wichtige Informationen zu Strategien, Wachstumsprognosen und Risiken, die anfänglich möglicherweise weniger negativ interpretiert oder erst verzögert wahrgenommen werden. Wenn Investoren die langfristigen Implikationen im Zeitverlauf gründlicher bewerten und erkennen, können negative Kursentwicklungen in der Driftperiode eintreten.

Zweitens könnte die unterschiedliche Behandlung der Informationen aus der Q&A-Sitzung und der Präsentation auf unterschiedliche Wahrnehmungen der Glaubwürdigkeit hinweisen. Q&A-Aussagen wirken oft direkter und authentischer, was sie für Investoren glaubwürdiger macht. Da sie als weniger manipulierbar angesehen werden, werden diese Informationen in der Driftperiode nicht korrigiert. Im Gegensatz dazu wird die Präsentation als strategisch und marketingorientiert betrachtet. Diese Wahrnehmung kann dazu führen, dass anfänglich positive Reaktionen auf die Präsentation in der Driftperiode relativiert werden, wenn Investoren die langfristige Tragfähigkeit und möglichen Überoptimismus in Frage stellen. Diese Korrektur zeigt sich in negativen Renditen, was darauf hinweist, dass das anfängliche Vertrauen in die Präsentation nachlässt, sobald mehr Zeit für die Analyse und Bewertung vergangen ist.

4.2.5 Vergleich der verwendeten Sentiment-Lexika

Die Unterschiede in den Ergebnissen, die sich aus der Anwendung verschiedener Lexika ergeben, verdeutlichen, dass die Wahl des Lexikons die Resultate hinsichtlich der Marktreaktionen und Interpretation der abnormalen Renditen stark beeinflusst und bestätigen die Hypothese 3a. Zusammengefasst scheinen das LM- und Henry-Lexikon besser zur Erfassung des Sentiments über den gesamten Verlauf eines Earnings Calls geeignet als das Harvard-GI Lexikon, da sie konsistentere und signifikante Ergebnisse liefern. Harvard-GI zeigt in allen Modellen die geringste Modellgüte und weist teilweise Unterschiede in den signifikanten Variablen auf, was darauf hindeutet, dass es

möglicherweise weniger gut für die spezifischen Inhalte der Earnings Calls geeignet ist oder Sentiment unpräziser misst.

Die Hypothese H3b, dass das mit dem LM-Lexikon ermittelte Sentiment einen ausgeprägteren Zusammenhang zur Marktreaktion zeigt als das Sentiment, das mit den Lexika von Henry oder Harvard GI ermittelt wird, kann in dieser Form nicht bestätigt werden. Die divergierende Modellgüte der Lexika in den verschiedenen Abschnitten des Earnings Calls demonstriert, dass kein einzelnes Lexikon als universell überlegen erachtet werden kann. Die Eignung eines Lexikons ist maßgeblich von der Textart und dem spezifischen Kontext abhängig. Da die Lexika für unterschiedliche Anwendungsbereiche entwickelt wurden, variiert ihre Fähigkeit zur Sentimentfassung in Abhängigkeit von dem jeweiligen Abschnitt des Earnings Calls. Das Harvard-GI-Lexikon wurde ursprünglich für die Sozialpsychologie erstellt und basiert auf allgemeinen psychosozialen Wortlisten.¹⁷¹ In Anbetracht der nicht auf den finanziellen Kontext ausgerichteten Konzeption des Lexikons erweist sich dieses als weniger geeignet für den Präsentationsteil von Earnings Calls, in dem eine spezifische Finanzsprache zum Einsatz kommt. Das Henry-Lexikon wurde hingegen speziell für die Analyse von Gewinnmitteilungen entwickelt¹⁷² und empfiehlt sich insbesondere für die Verwendung in Präsentationen, da diese in der Regel eine ähnliche Sprache verwenden. Auch das LM-Lexikon, das auf einer Analyse von US-Jahresberichten basiert,¹⁷³ ist in der Lage, derartige Präsentationen gut zu erfassen. Allerdings könnte die Präsentation eher einer Gewinnmitteilung als einem Jahresabschluss ähneln, was das LM-Lexikon im Vergleich zum Henry-Lexikon etwas benachteiligt.

Innerhalb des Q&As lassen sich die Aussagen der Unternehmensvertreter mit dem GI-Lexikon am adäquatesten erfassen, da es eine allgemeinere Sprache und eine höhere Affinität zu psychosozialen Inhalten aufweist. Dies ist insbesondere in einem spontanen Kontext von Vorteil. Im Vergleich zu den Formulierungen in Gewinnmitteilungen und Jahresberichten ist die von den Managern in diesem Abschnitt verwendete Sprache dynamischer und konversationsartiger, was die Anwendung der anderen beiden Lexika erschweren könnte. Die von Analysten vorbereiteten und spezifischen finanziellen Fragen lassen sich mittels finanzspezifischer Lexika besser analysieren als durch das Harvard GI-Lexikon. Infolge der umfangreicheren Wortanzahl ist davon auszugehen, dass das LM-Lexikon in diesem Abschnitt eine höhere Anpassungsfähigkeit und präzisere Messung aufweist als das Henry-Lexikon.

Die Weiterentwicklung und Anpassung der Lexika für die verschiedenen Abschnitte von Earnings Calls stellt eine vielversprechende Möglichkeit für eine Verbesserung der Sentiment-Erfassung dar. Eine Kombination mehrerer Lexika und eine Analyse jedes Abschnitts, mit dem am besten für diesen geeigneten Lexikon könnten zu einem umfassenderen und präziseren Verständnis der Stimmungslage führen.

¹⁷¹Vgl. Stone und Hunt (1963, pp. 241f).

¹⁷²Vgl. Henry (2008, p. 384).

¹⁷³Vgl. Loughran und McDonald (2016, p. 1204).

Zur Überprüfung dieser Annahme wird eine Regression durchgeführt, in der das Sentiment der Präsentation mit dem Henry-Lexikon, das Sentiment der Analystenfragen mit dem LM-Lexikon sowie das Sentiment der Antworten der Unternehmensleitung im Q&A des Earnings Calls mit dem Harvard GI-Lexikon ermittelt wird. Das Resultat belegt eine höhere Modellgüte im Vergleich zu allen vorherigen Modellen und signifikante Werte für Präsentation und Analystenfragen auf dem 0,1-%-Niveau sowie für Managementantworten auf dem 5-%-Niveau (siehe Tabelle 5). Während Hypothese H2c in der Betrachtung der Segmente mit einem einzelnen Lexikon nicht vollständig bestätigt werden konnte, belegt die Untersuchung mit kombinierten Lexika nun einen signifikanten und unabhängigen Einfluss des Sentiments der Manager und Analysten innerhalb des Q&A-Abschnitts auf die Marktreaktion. Zudem übertrifft dieses Modell hinsichtlich der Erklärungskraft sämtliche anderen Kombinationen der drei Lexika.

Die dargelegten Befunde verdeutlichen die Einschränkungen lexikonbasierter Sentimentmethoden, insbesondere im Hinblick auf die Erfassung von Kontext und Mehrdeutigkeit von Wörtern. Dabei weist die Quantifizierung der qualitativen Anlegerstimmung in einem Sentimentmaß inhärente Ungenauigkeiten auf.¹⁷⁴ Klassifikationsfehler oder die Nichterkennung spezifischer Begriffe können dabei zu einer Verzerrung der Sentimentmessung und somit zu einer Beeinträchtigung der Genauigkeit der geschätzten Regressionskoeffizienten führen.¹⁷⁵ Besonders in Earnings Calls, die verschiedene Abschnitte und Redner mit unterschiedlichen Themen und Stilen umfassen, erschwert dies eine präzise Sentimentanalyse und kann zu weniger verlässlichen Ergebnissen führen. Die Ungenauigkeiten in der Messung könnten die teilweise geringe Aussagekraft der Sentimentindikatoren erklären, ohne dass diese die tatsächliche Relevanz des Sentiments für Anleger reflektiert.

Angesichts dieser Einschränkungen lexikonbasierter Methoden könnte der Einsatz maschineller Lernverfahren eine vielversprechende Alternative darstellen, um die Sentimentanalyse in Earnings Calls zu präzisieren. Frankel et al. (2021) stellten fest, dass besonders in spontaneren und dynamischeren Textquellen die Vorteile maschinellen Lernens gegenüber lexikonbasierten Methoden deutlich werden.¹⁷⁶ Die Kenntnis spezifischer Bedeutungen und Tonfälle in unterschiedlichen Kontexten, wie verschiedenen Segmenten eines Earnings Calls, kann durch entsprechende Trainingsdaten erworben werden. In diesem Zusammenhang kann beispielsweise die thematische und stilistische Variation im Q&A-Abschnitt im Vergleich zur Präsentation bei der Klassifikation von Wörtern Berücksichtigung finden. Ein weiterer Vorteil des maschinellen Lernens besteht in der Fähigkeit, spezifische Begriffe, die in einem Lexikon möglicherweise fehlen, durch Modelltraining zu identifizieren

und zu verarbeiten.

4.3 Einfluss des Sentiments in Abhängigkeit von der Informationslage

Die Reaktion der Investoren auf den Tonfall in Earnings Calls dürfte maßgeblich durch ihre jeweilige Informationslage beeinflusst werden. Dabei spielen sowohl persönliche Anlegerfaktoren als auch das Informationsumfeld des Unternehmens eine Rolle. In dieser Arbeit werden daher der Anteil institutioneller Anleger, die S&P 500-Zugehörigkeit, der ESG-Score und die Dividendenausschüttung untersucht. Da für die genannten Aspekte keine quartalsweisen Daten verfügbar sind oder die Werte über die Zeit hinweg stabil bleiben, können sie nicht direkt in die Regression einbezogen werden. Dies ist insbesondere relevant für Panelregressionen mit fixen Effekten, die Wertänderungen berücksichtigen. Das Ergebnis eines Hausman-Tests weist einen p-Wert nahe Null auf, was darauf hinweist, dass eine Random-Effects-Schätzung in diesem Fall inkonsistent wäre und das Fixed-Effects-Modell zu bevorzugen ist. Um dennoch den Einfluss des Anteils institutioneller Investoren und des ESG-Scores zu analysieren, wurden die Daten nach dem Terzil dieser Variablen unterteilt und separate Regressionen für das höchste und das niedrigste Terzil durchgeführt. Für die Zugehörigkeit zum S&P 500 und das Zahlen von Dividenden werden die Beobachtungen in Gruppen eingeteilt und beide Gruppen direkt miteinander verglichen. Für die Analyse wird jeweils das Sentiment in der Präsentation mit dem Henry-Lexikon, das der Manager Aussagen im Q&A mit dem Harvard-GI-Lexikon und das der Analystenfragen mit dem LM-Lexikon ermittelt.

4.3.1 Anteil institutioneller Anleger

Zunächst wird der Zusammenhang zwischen der Anlegerkompetenz, gemessen am Anteil institutioneller Investoren, und der Marktbewertung des linguistischen Tons in Earnings Calls untersucht. Ein hoher Anteil institutioneller Anleger lässt auf eine Beobachtung des Unternehmens durch erfahrene und fachkundige Marktteilnehmer mit umfassenden Informationsressourcen und Expertise schließen.¹⁷⁷ In Übereinstimmung mit den Studien von Bochkay et al. (2020) und Brockman et al. (2015) zeigen die Ergebnisse in Tabelle 7, dass Unternehmen mit einem höheren Anteil institutioneller Investoren stärkere Marktreaktionen auf das Sentiment in allen Abschnitten der Earnings Calls verzeichnen. Für institutionelle Anleger ist das Sentiment stärker signifikant, weist höhere Koeffizienten auf und zeigt eine bessere Modellgüte. Während für private Anleger vor allem die Präsentation von Relevanz ist, legen institutionelle Investoren großen Wert auf die Stimmung der Unternehmensvertreter im Q&A. Dies deutet darauf hin, dass institutionelle Investoren ihre Entscheidungen auf eine umfassendere und detailliertere Analyse der erhaltenen Informationen stützen, während private Anleger stärker auf einfach zugängliche und klar strukturierte Inhalte angewiesen sind. Die Ergebnisse legen insgesamt nahe, dass

¹⁷⁴Vgl. Loughran und McDonald (2016, pp. 1189–1191).

¹⁷⁵Vgl. Frankel et al. (2021, p. 5516); Loughran und McDonald (2016, p. 1226).

¹⁷⁶Vgl. Frankel et al. (2021, pp. 5516–5518).

¹⁷⁷Vgl. Bochkay et al. (2019, p. 50); Brockman et al. (2015, pp. 25 und 28).

Tabelle 7: Einfluss des Sentiments in Earnings Calls auf die CAR in Abhängigkeit vom Anteil institutioneller Anleger (IO)

	CAR(-1,1)		CAR(2,42)	
	IO gering	IO hoch	IO gering	IO hoch
Präsentation	0.068** (0.0255)	0.088*** (0.0179)	-0.103** (0.0386)	-0.033 (0.0310)
Manager Q&A	0.084 (0.0459)	0.108** (0.0407)	0.034 (0.0731)	-0.036 (0.0580)
Analyst Q&A	0.047*** (0.0113)	0.048*** (0.0108)	-0.025 (0.0200)	0.007 (0.0151)
R ²	0.0931	0.1037	0.1347	0.1462
Adjusted R ²	-0.0035	0.0245	0.0425	0.0708

*** $p < 0.1\%$ ** $p < 1\%$ * $p < 5\%$. $p < 10\%$

das Sentiment für institutionelle Investoren bedeutsamer ist und sie besser in der Lage sind, wertrelevante Informationen aus dem linguistischen Ton zu extrahieren als weniger geschulte individuelle Investoren. Letztere sind stärker auf offensichtliche, quantitative Informationen angewiesen und zeigen eine geringere Fähigkeit, subtile Nuancen im Tonfall der Teilnehmer des Earnings Calls zu interpretieren.

Diese Beobachtung zeigt sich auch in der Driftperiode: Die negative Signifikanz des Präsentationsentiments in der Driftperiode bei privaten Anlegern deutet darauf hin, dass sie diese Signale erst verzögert wahrnehmen oder anfangs weniger negativ interpretieren, was zu späteren negativen Kurskorrekturen führt. Bei Unternehmen mit einem hohen Anteil institutioneller Investoren zeigt das Präsentationsentiment in der Driftperiode hingegen keinen signifikanten Einfluss. Institutionelle Investoren bewerten die Informationen demnach bereits im initialen Reaktionsfenster genauer. Die Aussagen von Analysten und Managern während des Q&A-Abschnitts sind für beide Gruppen in der Driftperiode nicht signifikant.

4.3.2 S&P500 Zugehörigkeit

Zweitens zeigt ein Vergleich zwischen Unternehmen, die im S&P 500 gelistet sind, und solchen, die dies nicht sind, dass Investoren kleinerer Unternehmen empfindlicher auf Kommunikationsaspekte in Earnings Calls reagieren (siehe Tabelle 8). Die strengen regulatorischen Anforderungen sowie die starke öffentliche Aufmerksamkeit führen bei S&P-500-Unternehmen zu einer umfassenderen Informationsverfügbarkeit.¹⁷⁸ Das Sentiment in Präsentationen und Analystenfragen wirkt sich signifikant auf die abnormalen Renditen beider Gruppen aus, wobei dieser Effekt bei Unternehmen außerhalb des S&P 500 stärker ausgeprägt ist. Des Weiteren lässt sich ein signifikanter Einfluss der Aussagen von Managern während der Q&A-Sitzungen lediglich bei diesen Unternehmen feststellen. Diese Erkenntnis lässt sich durch höhere Unsicherheit und geringere Informationsverfügbarkeit erklären, weshalb Anleger stärker auf zusätzliche Informationen und Nuancen in spontanen Antworten des

Managements angewiesen sind. Die Beobachtungen stehen im Einklang mit der Annahme von Milian et al. (2017), die den vergleichsweise geringen ökonomischen Einfluss des Sentiments von S&P-500-Unternehmen in ihrer Studie auf deren effiziente Bewertung zurückführen.¹⁷⁹

In der Driftperiode lässt sich ein signifikant negativer Einfluss des Sentiments der Präsentation nur bei Unternehmen beobachten, die nicht dem S&P 500 angehören. Dies deutet darauf hin, dass diese Unternehmen einer höheren Anfälligkeit für nachträgliche Neubewertungen sowie einer verzögerten Erkennung langfristiger Risiken unterliegen, während bei S&P-500-Unternehmen unmittelbar eine bessere Verfügbarkeit und gründlichere Bewertung der Informationen vorliegt. Die Aussagen von Analysten und Managern während der Q&A-Sitzungen üben in der Driftperiode für Investoren beider Gruppen keinen signifikanten Einfluss auf die abnormalen Renditen aus.

4.3.3 ESG-Score

Drittens wird der Zusammenhang zwischen der Nachhaltigkeit eines Unternehmens, bewertet anhand seines ESG-Scores, und dem Einfluss des Sentiments in Earnings Calls untersucht. Der von LSEG ermittelte ESG-Score basiert auf vom Unternehmen selbst bereitgestellten Informationen zu Umwelt-, Sozial- und Unternehmensführungsaspekten und dient der Messung der relativen Nachhaltigkeit der ESG-Berichterstattung eines Unternehmens.¹⁸⁰ Unternehmen, die hohe ESG-Scores aufweisen, zeichnen sich durch eine gesteigerte Transparenz sowie ethischere Geschäftspraktiken aus. Dies führt zu einer Verringerung der Informationsasymmetrie zwischen Unternehmen und Investoren sowie der Agency-Probleme, da Investoren Zugang zu umfassenderen und verlässlicheren Informationen erhalten.¹⁸¹ Die Ergebnisse der Regressionen in Tabelle 9 legen nahe, dass Investoren bei Unternehmen mit niedrigem ESG-Score insgesamt stärker auf das Sentiment reagieren. Eine differenzierte Betrachtung der Abschnitte in den Earnings Calls zeigt, dass diese je nach ESG-Score unterschiedlich gewichtet

¹⁷⁹Vgl. Milian et al. (2017, p. 73).

¹⁸⁰Vgl. LSEG (2024).

¹⁸¹Vgl. Huang et al. (2024), p. 1156.

¹⁷⁸Vgl. H. Chen et al. (2004, p. 1928).

Tabelle 8: Einfluss des Sentiments in Earnings Calls auf die CAR in Abhängigkeit von der Zugehörigkeit zum S&P 500 Index

	CAR(-1,1)		CAR(2,42)	
	S&P500	Rest	S&P500	Rest
Präsentation	0.075*** (0.0162)	0.085*** (0.0158)	-0.029 (0.0304)	-0.097*** (0.0243)
Manager Q&A	0.003 (0.0324)	0.080* (0.0318)	-0.011 (0.0629)	-0.013 (0.0452)
Analyst Q&A	0.030*** (0.0085)	0.056*** (0.0083)	0.003 (0.0133)	-0.009 (0.0130)
R ²	0.0683	0.1106	0.1228	0.1286
Adjusted R ²	-0.0244	0.0412	0.0356	0.0606

*** p < 0.1% ** p < 1% * p < 5% . p < 10%

Tabelle 9: Einfluss des Sentiments in Earnings Calls auf CAR in Abhängigkeit vom ESG-Score

	CAR(-1,1)		CAR(2,42)	
	ESG gering	ESG hoch	ESG gering	ESG hoch
Präsentation	0.084*** (0.0218)	0.111*** (0.0183)	-0.131*** (0.0350)	-0.071** (0.0272)
Manager Q&A	0.100* (0.0477)	0.033 (0.0373)	0.116 (0.0763)	-0.038 (0.0593)
Analyst Q&A	0.050*** (0.0121)	0.043*** (0.0110)	-0.020 (0.0191)	0.010 (0.0159)
R ²	0.1112	0.0939	0.1366	0.1300
Adjusted R ²	0.0251	0.0069	0.0529	0.0464

*** p < 0.1% ** p < 1% * p < 5% . p < 10%

werden. Bei Unternehmen mit niedrigem ESG-Score und geringerer Transparenz müssen Investoren alle verfügbaren Informationen, einschließlich der spontanen Aussagen im Q&A, intensiv nutzen, um Unsicherheiten zu reduzieren und fundierte Entscheidungen zu treffen. Unternehmen mit einem hohen ESG-Score genießen demgegenüber ein größeres Vertrauen, sodass Investoren sich in stärkerem Maße auf das Sentiment in den vorbereiteten Präsentationen verlassen. Die spontane Frage- und Antwortrunde besitzt in diesem Kontext keinen signifikanten Einfluss, da anzunehmen ist, dass die vorab bereitgestellten Informationen von den Investoren als verlässlich und ausreichend für fundierte Entscheidungen eingeschätzt werden. Unabhängig vom ESG-Score spielen Analystenfragen eine zentrale Rolle für Investoren aller Unternehmen.

Die Untersuchung des Einflusses des ESG-Scores auf die Investorenreaktionen während der Driftperiode ergibt, dass das Sentiment der Präsentation bei Unternehmen mit niedrigem ESG-Score einen signifikanteren und stärker negativen Einfluss auf die abnormalen Renditen hat als bei Unternehmen mit hohem ESG-Score. Die geringere Transparenz und die höheren Risiken bei diesen Unternehmen bedingen eine intensivere, nachgelagerte Analyse durch Investoren, was zu stärkeren Anpassungen im zeitlichen Verlauf führt. Bei nachhaltigeren Unternehmen besteht für Investoren aufgrund der umfassenderen und verlässlicheren Informationen offenbar ein geringerer Bedarf, ihre anfänglichen Einschätzungen in der Driftperiode zu korrigieren. Aussagen von Ana-

lysten und Managern während der Q&A-Sitzungen erweisen sich demgegenüber für beide Gruppen in der Driftperiode als nicht signifikant.

4.3.4 Dividendenzahlungen

Abschließend stellen Dividenden einen wichtigen Indikator für die finanzielle Stabilität und die Sicherheit zukünftiger Cashflows eines Unternehmens dar.¹⁸² Die Sentimentvariablen weisen in der initialen Reaktion bei beiden Gruppen vergleichbare Signifikanzniveaus auf. In Übereinstimmung mit Price et al. (2012) kann anhand des Vergleichs der Koeffizienten festgestellt werden, dass die Reaktion der Investoren auf die verschiedenen Aspekte des Sentiments in Earnings Calls bei Unternehmen, die Dividenden zahlen, etwas geringer ausfällt (siehe Tabelle 10). Die positive Signalwirkung der Dividenden kann demnach dazu führen, dass Investoren bei diesen Unternehmen anfänglich das Sentiment nicht umfassend berücksichtigen.

Zusammenfassend legen die Ergebnisse nahe, dass eine geringere Informationslage der Investoren eines Unternehmens zu einer stärkeren Reaktion auf das Sentiment in seinen Earnings Calls führt, was die Hypothese 1b bestätigt. Die Kompetenz der Anleger sowie das Informationsumfeld des Unternehmens spielen eine zentrale Rolle bei der Formulierung ihrer Erwartungen und der Verarbeitung von Informa-

¹⁸²Vgl. Price et al. (2012, pp. 995f).

Tabelle 10: Einfluss des Sentiments in Earnings Calls auf die CAR in Abhängigkeit von dem Zahlen von Dividenden

	CAR(-1,1)		CAR(2,42)	
	Dividenden	Keine	Dividenden	Keine
Präsentation	0.072*** (0.0138)	0.107*** (0.0230)	-0.079*** (0.0219)	-0.0634 (0.0392)
Manager Q&A	0.050 (0.0281)	0.082 (0.0437)	-0.048 (0.0398)	0.028 (0.0717)
Analyst Q&A	0.044*** (0.0080)	0.057*** (0.0113)	0.000 (0.0107)	-0.014 (0.0194)
R ²	0.0884	0.1128	0.1599	0.1076
Adjusted R ²	0.0159	0.0311	0.0931	0.0254

*** p < 0.1% ** p < 1% * p < 5% . p < 10%

tionen. Unternehmen, die einen geringen Anteil institutioneller Investoren aufweisen, nicht im S&P 500 gelistet sind, einen niedrigen ESG-Score besitzen oder keine Dividenden ausschütten, erfahren eine stärkere Reaktion der Investoren auf qualitative Informationen aus dem Tonfall in Earnings Calls. Insbesondere bei diesen Unternehmen kann die Auswertung von Texten folglich einen wertvollen Beitrag zur Finanzanalyse leisten. Für weniger geschulte individuelle Investoren stellt die Interpretation subtiler Nuancen im Tonfall der Teilnehmer des Earnings Calls eine Herausforderung dar.

Darüber hinaus lässt sich festhalten, dass auch die Relevanz der unterschiedlichen Segmente von Earnings Calls maßgeblich von der Informationslage des jeweiligen Unternehmens abhängig ist. Die Präsentation stellt im Vergleich zu den einzelnen Bestandteilen des Q&As den größten und unmittelbarsten Einfluss auf die Entscheidungsfindung dar. Für Unternehmen, die durch eine höhere Unsicherheit und geringere Transparenz gekennzeichnet sind, erlangen die spontanen Antworten des Managements im Rahmen des Q&As eine entscheidendere Bedeutung. Diese bieten den Investoren zusätzlichen Informationswert, indem sie einen Einblick in unvorhergesehene Themen und die Fähigkeit des Managements geben, spontan auf wichtige Fragen zu reagieren. Während institutionelle Investoren ihre Entscheidungen auf eine tiefere Analyse des Q&As stützen, reagieren private Anleger stärker auf die leichter zugänglichen Inhalte der Präsentation. Unabhängig von den spezifischen Unternehmensmerkmalen liefern die Analysten in dem initialen Ereignisfenster für Investoren stets wertvolle Informationen, wobei ihr Einfluss bei Unternehmen mit geringerer Informationsverfügbarkeit tendenziell stärker ausfällt. Die anfängliche Marktreaktion auf die Stimmung der Unternehmensvertreter in der Präsentation wird verstärkt bei Unternehmen mit unzureichender Informationslage korrigiert, während bei informationseffizienten Unternehmen weniger Anpassungen erforderlich sind. Insbesondere private Anleger erfassen Signale dabei oft verzögert, was zu späteren Korrekturen führt.

4.4 Robustheit

4.4.1 Alternative Ermittlung der Gewinnüberraschung

Die fehlende Signifikanz der Gewinnüberraschung in der initialen Marktreaktion ist insofern überraschend, als sie in

anderen Veröffentlichungen häufig als maßgeblicher quantitativer Einflussfaktor auf abnormale Renditen identifiziert wurde.¹⁸³ Um diese Beobachtung zu untersuchen, wird zusätzlich ein weiterer etablierter Ansatz zur Messung der Gewinnüberraschung herangezogen. Im Rahmen dieses Ansatzes erfolgt ein Vergleich der tatsächlich gemeldeten Gewinne $EPS_{i,q}$ pro Aktie eines Unternehmens i im Quartal q mit den von Analysten antizipierten Gewinnen $AnalystEPS_{i,q}$:

$$AnalystSURP_{i,q} = \frac{EPS_{i,q} - AnalystEPS_{i,q}}{AnalystEPS_{i,q}} * 100. \quad (11)$$

Auch die anhand von Analystenprognosen ermittelte Gewinnüberraschung zeigt im dreitägigen Ereignisfenster um den Earnings Call keinen signifikanten Einfluss und ist im Gegensatz zur saisonalen Messung auch in der Driftperiode nicht signifikant. Des Weiteren fällt die Modellgüte in beiden Zeiträumen leicht geringer aus als bei der saisonalen Messung. Die fehlende Signifikanz der Gewinnüberraschung in der initialen Marktreaktion, gemessen durch den saisonalen und den Analystenansatz, lässt den Schluss zu, dass in dem spezifischen Kontext dieser Arbeit andere Faktoren, wie das Sentiment der Teilnehmer oder externe Einflüsse, eine wichtigere Rolle spielen. In Bezug auf das Vorzeichen und die Signifikanz der Sentimentvariablen lassen sich bei beiden Ansätzen zur Ermittlung der Gewinnüberraschung konsistente Ergebnisse beobachten, was deren Robustheit unterstreicht.

4.4.2 Alternative Ermittlung abnormaler Renditen

Wie zuvor in der Methodik dargelegt, können unterschiedliche Methoden zur Ermittlung abnormaler Renditen zu stark unterschiedlichen Ergebnissen führen.¹⁸⁴ Zur Prüfung der Robustheit der Ergebnisse wird daher neben dem Marktmodell auch ein faktorloses Modell analysiert, das in einer Vielzahl von Studien zum Sentiment in Earnings Calls Anwendung findet.¹⁸⁵ In faktorlosen Modellen

¹⁸³Vgl. De Amicis et al. (2021); Milian et al. (2017, p. 70); Price et al. (2012, p. 998).

¹⁸⁴Vgl. Gehrke (2022, p. 279).

¹⁸⁵Vgl. Mayew et al. (2020, p. 316); Bochkay et al. (2019, p. 36); Milian et al. (2017, p. 70); Brockman et al. (2015, p. 26); Lee (2015, p. 235); Price et al. (2012, p. 998).

erfolgt die Ermittlung der erwarteten Rendite auf Basis festgelegter Benchmarks, wie beispielsweise anhand eines Vergleichsportfolios anderer Wertpapiere.¹⁸⁶ In dieser Arbeit wird die erwartete Rendite eines Unternehmens i am Tag t $E(R_{i,t})$ als mittlere Rendite $\overline{R_{p,t}}$ eines Portfolios von Aktien mit dem gleichen Größen- und Buchwert-zu-Markt-Quartil wie Unternehmen i berechnet:

$$E(R_{i,t}) = \overline{R_{p,t}}. \quad (12)$$

Im Vergleich zum Marktmodell zeigt das faktorlose Modell im initialen Ereignisfenster eine leicht geringere Modellgüte, in der Driftperiode hingegen einen wesentlich geringeren Wert. Das deutet darauf hin, dass das Marktmodell insbesondere die längerfristige Entwicklung nach den Earnings Calls besser erfasst. Die nahezu identischen Sentiment-Koeffizienten in beiden Modellen in der initialen Reaktion weisen sowohl auf die Robustheit der Ergebnisse als auch auf eine konsistente Beziehung zwischen erklärenden Variablen und abnormalen Renditen hin. In der Driftperiode zeigen sich jedoch Unterschiede: Während die Sentimentvariablen des faktorlosen Modells keinen signifikanten Einfluss haben, weist das Präsentations-sentiment im Marktmodell einen signifikant negativen Effekt auf. Diese Beobachtung deutet darauf hin, dass das faktorlose Modell die langfristige Beziehung zwischen dem Sentiment und den abnormalen Renditen nicht zuverlässig erfasst, möglicherweise aufgrund des fehlenden Einbezugs systematischer Marktfaktoren, die in der Driftperiode eine größere Rolle spielen könnten.

Des Weiteren ist festzustellen, dass die Gewinnüberraschung in der initialen Reaktion lediglich im faktorlosen Modell signifikant ist. Dies lässt vermuten, dass das Modell die direkte Beziehung zwischen Gewinnüberraschungen und abnormalen Renditen direkter erfasst, da es weniger durch den allgemeinen Markttrend beeinflusst wird. Der Marktindex könnte die Signifikanz der Gewinnüberraschung abschwächen, indem er die allgemeine Marktstimmung einbezieht, die auch auf andere Informationen reagiert.

Insgesamt bestätigen die Ergebnisse die Robustheit der gefundenen Effekte und zeigen, dass unterschiedliche Ansätze zu der Ermittlung abnormaler Renditen zu variierenden, aber in den Kernbefunden konsistenten Ergebnissen führen.

5 Fazit

Die theoretische Betrachtung dieser Arbeit hat gezeigt, dass qualitative Informationen, insbesondere im Rahmen von Earnings Calls, in der Finanzanalyse eine zunehmend wichtige Rolle spielen. Diese Form der Unternehmenskommunikation stellt eine dynamische Plattform zur Verfügung, auf der über die reine Vermittlung von Zahlen hinausgehende, wertvolle emotionale Hinweise vermittelt werden. Die textbasierte Sentimentanalyse stellt dabei ein zentrales Instrument dar, um diese sprachlichen Nuancen systematisch zu

erfassen und zu interpretieren. Sie ermöglicht ein besseres Verständnis der Auswirkungen der Stimmung von Unternehmensvertretern und Analysten auf die Marktreaktionen und leistet damit einen wesentlichen Beitrag zur Vertiefung des Wissens über die Marktpsychologie und die Informationsverarbeitung von Investoren.

Basierend auf den theoretischen Erkenntnissen zur Rolle des Sentiments in der Finanzanalyse bestätigt die empirische Untersuchung, dass emotionale und sprachliche Nuancen in Unternehmensoffenlegungen entscheidende Signale für Marktteilnehmer darstellen. Investoren betrachten das Sentiment in Earnings Calls als wichtige Informationsquelle und integrieren es fortlaufend in ihre Entscheidungsprozesse. Nach einer anfänglich emotional geprägten Reaktion bewerten die Investoren zunehmend die langfristigen Implikationen der Informationen, was in der Driftperiode oft zu einer negativen Korrektur der anfänglich positiven Marktreaktionen führt. Die Ergebnisse legen dabei insgesamt nahe, dass eine differenzierte Herangehensweise an die Sentimentanalyse notwendig ist, um die komplexen Zusammenhänge zwischen den verschiedenen Segmenten eines Earnings Calls und den damit verbundenen Marktreaktionen umfassend zu erfassen. Das Sentiment in dem Q&A-Abschnitt ist für die initialen Marktreaktionen bedeutender als die Präsentation. Innerhalb des Q&As zeigen die Fragen der Analysten zudem einen größeren Einfluss auf die Marktreaktionen als die Antworten des Managements. Während die Präsentation im zeitlichen Verlauf an Bedeutung gewinnt und zu einer Revision der anfänglich positiven Reaktion führt, bleiben die authentischeren Informationen aus dem Q&A-Abschnitt stabiler und weniger anfällig für Korrekturen.

Ferner wurde deutlich, dass die Reaktionen der Investoren auf das Sentiment in Earnings Calls in hohem Maße von ihrer Informationslage abhängig sind. Unternehmen, die einen geringen Anteil institutioneller Investoren aufweisen, nicht dem S&P 500 angehören, niedrige ESG-Scores besitzen oder keine Dividenden ausschütten, erfahren intensivere Marktreaktionen auf den Tonfall in den Earnings Calls. Im Verlauf der Driftperiode wird die anfänglich positive Marktreaktion in diesen Fällen zudem häufiger revidiert. Besonders bei Unternehmen mit höherer Unsicherheit und begrenztem Informationszugang bietet eine Analyse des Sentiments somit wertvolle Einblicke. Erfahrene institutionelle Anleger sind in der Lage, den linguistischen Ton in Earnings Calls präziser zu interpretieren und daraus gezielt Informationen für ihre Investitionsentscheidungen zu extrahieren. Auch die Bedeutung der einzelnen Segmente von Earnings Calls variiert mit der Informationslage: Die spontanen Antworten des Managements im Q&A gewinnen bei höheren Informationsasymmetrien gegenüber der Präsentation an Gewicht. Zudem scheinen professionelle Anleger besonders gut in der Lage zu sein, wertvolle Informationen aus dem Manager-Sentiment in dem Q&A zu extrahieren, während weniger erfahrene Investoren stärker auf die Präsentation fokussiert sind. Unabhängig von unternehmensbezogenen Merkmalen oder der Erfahrung der Investoren ist die Stimmung der Analysten stets von entscheidender Bedeutung.

¹⁸⁶Vgl. Gerpott (2009, p. 216).

Darüber hinaus wurde ein besonderer Fokus auf die methodischen Herausforderungen einer lexikonbasierten Sentimentanalyse gelegt, um die Potenziale und Grenzen dieser Methode im Bereich der Finanzanalyse besser zu verstehen. Die lexikonbasierte Sentimentanalyse hat sich als nützlich erwiesen, allerdings zeigten sich Einschränkungen bei der Erfassung von Kontext und Mehrdeutigkeit, was die Genauigkeit der Sentiment-Messung beeinträchtigt. Insbesondere die Wahl des Lexikons ist zentral für präzise Ergebnisse, wobei keines der in dieser Arbeit verwendeten Lexika universell überlegen war. Stattdessen hängt die Eignung stark vom spezifischen Abschnitt des Earnings Calls ab. Die Arbeit legt nahe, dass eine kontextabhängige Auswahl und Kombination von Lexika eine Verbesserung der Ergebnisse bieten kann. Dabei verdeutlichen die aufgezeigten Herausforderungen, dass der Einsatz maschineller Lernverfahren einen vielversprechenden Ansatz zur Verbesserung der Präzision der Sentimentanalyse in Earnings Calls darstellen könnte.

Insgesamt liefert die Arbeit somit wichtige Erkenntnisse darüber, wie qualitative Informationen aus Earnings Calls systematisch erfasst und für die Finanzanalyse genutzt werden können. Allerdings bedarf die komplexe Dynamik zwischen den verschiedenen Informationsquellen und Analysemethoden weiterer Untersuchung. Auf Grundlage der Ergebnisse lassen sich vier zentrale Ansatzpunkte für zukünftige Forschungen identifizieren: Erstens sollte die Weiterentwicklung von Sentiment-Analysemethoden im Vordergrund stehen, wobei die Integration maschinellen Lernens vielversprechende Möglichkeiten bietet, um Stimmungen in den unterschiedlichen Segmenten der Earnings Calls präziser zu erfassen. Zweitens reduziert die Analyse von textuellen Transkripten im Vergleich zu Videoübertragungen zwar die Komplexität, kann jedoch auch mit einem möglichen Informationsverlust verbunden sein. Zukünftige Untersuchungen sollten daher auch erforschen, wie nonverbale Faktoren wie Mimik, Gestik und Körpersprache in die Sentimentanalyse einbezogen werden können.

Drittens sollte sich zukünftige Forschung intensiver mit dem Einfluss der Informationslage der Investoren auf die Interpretation und Reaktion der Investoren auf das Sentiment in Earnings Calls befassen. In diesem Zusammenhang scheint es empfehlenswert, sowohl das Informationsumfeld des Unternehmens, beispielsweise durch eine Segmentierung nach Branchen oder Unternehmen in stark regulierten Märkten, als auch persönliche Anlegerfaktoren wie die geografische Herkunft der Investoren zu untersuchen. Viertens erfordert die Sentimentanalyse eine differenzierte Herangehensweise, um die Zusammenhänge zwischen den Abschnitten und Aussagen verschiedener Redner in Earnings Calls sowie die daraus resultierenden Marktreaktionen besser zu verstehen. Insbesondere die Aussagen von CEOs und CFOs sollten inhaltlich und methodisch, etwa durch den Einsatz spezifischer Lexika, differenziert analysiert werden. Dies könnte auch den Einfluss von Agency-Problemen und persönlichen Merkmalen der Sprecher tiefer beleuchten und somit aufzeigen, wie unterschiedliche Perspektiven und potenzielle Interessenkonflikte die Wahrnehmung der

Investoren beeinflussen.

In einer zunehmend datengetriebenen Finanzwelt bleibt es entscheidend, die komplexen Wechselwirkungen zwischen qualitativen Informationen und Marktreaktionen weiter zu erforschen. Die Fähigkeit, Stimmungen und psychologische Signale in Unternehmensveröffentlichungen wie Earnings Calls korrekt zu deuten, könnte für Investoren einen entscheidenden Vorteil bei der Unternehmensbewertung und der Vorhersage von Kursbewegungen darstellen. Damit wird die präzise Auswertung von Texten zu einem unverzichtbaren Bestandteil der modernen Finanzanalyse.

Literatur

- Agrawal, S., Azar, P. D., Lo, A. W., & Rai, T. S. (2018). Momentum, Mean-Reversion, and Social Media: Evidence from StockTwits and Twitter. *The Journal of Portfolio Management*, 44(7), 85–95.
- Ajinkya, B. B., & Gift, M. J. (1984). Corporate Managers' Earnings Forecasts and Symmetrical Adjustments of Market Expectations. *Journal of Accounting Research*, 22(2), 425–444.
- Allee, K. D., & Deangelis, M. D. (2015). The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion. *Journal of Accounting Research*, 53(2), 241–274.
- Altman, A. M. (2019). Petition for Rulemaking to Require Filers to Submit Earnings Conference Call Transcripts on Form 8-K, Pursuant to Regulation FD Compliance. Verfügbar 17. Mai 2024 unter <https://www.sec.gov/files/rules/petitions/2019/petn4-742.pdf>
- Antweiler, W., & Frank, M. Z. (2004). Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance*, 59(3), 1259–1294.
- Arslan-Ayaydin, Ö., Boudt, K., & Thewissen, J. (2016). Managers set the tone: Equity incentives and the tone of earnings press releases. *Journal of Banking & Finance*, 72, 132–147.
- Baddeley, M. (2018). *Behavioural Economics and Finance* (Second). Routledge.
- Bai, C., Duan, Y., Fan, X., & Tang, S. (2023). Financial market sentiment and stock return during the COVID-19 pandemic. *Finance Research Letters*, 54, 103738.
- Baker, H. K., & Nofsinger, J. R. (2010). *Behavioral Finance: Investors, Corporations, and Markets*. John Wiley & Sons.
- Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives*, 21(2), 129–151.
- Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 6(2), 159–178.
- Bassyouny, H., Abdelfattah, T., & Tao, L. (2022). Narrative disclosure tone: A review and areas for future research. *Journal of International Accounting, Auditing and Taxation*, 49, 100511.
- Beyer, A., Cohen, D. A., Lys, T. Z., & Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics*, 50(2-3), 296–343.
- Blau, B. M., DeLisle, J. R., & Price, S. M. (2015). Do sophisticated investors interpret earnings conference call tone differently than investors at large? Evidence from short sales. *Journal of Corporate Finance*, 31, 203–219.
- Bochkay, K., Hales, J., & Chava, S. (2019). Hyperbole or Reality? Investor Response to Extreme Language in Earnings Conference Calls. *The Accounting Review*, 95(2), 31–60.
- boerse.de. (n. d.). Was ist das Sentiment? Verfügbar 17. Mai 2024 unter <https://www.boerse.de/grundlagen/charttechnik/Was-ist-das-Sentiment>
- Borochin, P. A., Cicon, J. E., DeLisle, R. J., & Price, S. M. (2018). The effects of conference call tones on market perceptions of value uncertainty. *Journal of Financial Markets*, 40, 75–91.
- Bowen, R. M., Davis, A. K., & Matsumoto, D. A. (2002). Do conference calls affect analysts' forecasts? *The Accounting Review*, 77(2), 285–316.
- Brockman, P., Li, X., & Price, S. M. (2015). Differences in Conference Call Tones: Managers vs. Analysts. *Financial Analysts Journal*, 71(4), 24–42.

- CFI Education Inc. (2024). Earnings call. Verfügbar 10. Mai 2024 unter <https://corporatefinanceinstitute.com/resources/valuation/earnings-call/>
- Chen, H., Noronha, G., & Singal, V. (2004). The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation. *The Journal of Finance*, 59(4), 1901–1929.
- Chen, X., & Tian, Y. (2018). Learning to Perform Local Rewriting for Combinatorial Optimization. <https://doi.org/10.48550/arXiv.1810.00337>
- Chen, Y., Han, D., & Zhou, X. (2023). Mining the emotional information in the audio of earnings conference calls: A deep learning approach for sentiment analysis of securities analysts' follow-up behavior. *International Review of Financial Analysis*, 88, 102704.
- Cicon, J. (2015). Say it again Sam: the information content of corporate conference calls. *Review of Quantitative Finance and Accounting*, 48(1), 57–81.
- Collins, D. W., Nguyen, N. Q., & Nguyen, T. T. (2024). Manager sentiment and conditional conservatism. *Journal of Business Finance & Accounting*.
- Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web. *Management Science*, 53(9), 1375–1388.
- D'Augusta, C., & DeAngelis, M. D. (2019). Tone Concavity around Expected Earnings. *The Accounting Review*, 95(1), 133–164.
- Davis, A. K., Ge, W., Matsumoto, D., & Zhang, J. L. (2014). The effect of manager-specific optimism on the tone of earnings conference calls. *Review of Accounting Studies*, 20(2), 639–673.
- Davis, A. K., & Tama-Sweet, I. (2012). Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. *Contemporary Accounting Research*, 29(3), 804–837.
- De Amicis, C., Falconieri, S., & Tastan, M. (2021). Sentiment analysis and gender differences in earnings conference calls. *Journal of Corporate Finance*, 71, 101809.
- DeBoskey, D. G., Luo, Y., & Zhou, L. (2018). CEO power, board oversight, and earnings announcement tone. *Review of Quantitative Finance and Accounting*, 52(2), 657–680.
- Druz, M., Petzev, I., Wagner, A. F., & Zeckhauser, R. J. (2020). When Managers Change Their Tone, Analysts and Investors Change Their Tune. *Financial Analysts Journal*, 76(2), 47–69.
- Elliott, W. B., Loftus, S., & Winn, A. (2023). To read or to listen? Does disclosure delivery mode impact investors' reactions to managers' tone language? *Contemporary Accounting Research*, 41(1), 7–38.
- Feldman, R., Govindaraj, S., Livnat, J., & Segal, B. (2009). Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies*, 15(4), 915–953.
- Frankel, R., Jennings, J., & Lee, J. (2021). Disclosure Sentiment: Machine Learning vs. Dictionary Methods. *Management Science*, 68(7), 5514–5532.
- Frankel, R., Johnson, M., & Skinner, D. J. (1999). An Empirical Examination of Conference Calls as a Voluntary Disclosure Medium. *Journal of Accounting Research*, 37(1), 133–150.
- Fu, X., Wu, X., & Zhang, Z. (2019). The Information Role of Earnings Conference Call Tone: Evidence from Stock Price Crash Risk. *Journal of Business Ethics*, 173(3), 643–660.
- Gehrke, M. (2022). *Angewandte empirische Methoden in Finance & Accounting* (2. Aufl.). Springer Gabler.
- Gerpott, T. J. (2009). Ereignisstudie. In C. Baumgarth, M. Eisend & H. Evanschitzky (Hrsg.), *Empirische Mastertechniken. Eine anwendungsorientierte Einführung für die Marketing und Managementforschung* (S. 203–235). Gabler.
- Giancarlo, C. H. (2024). *Pure Storage, Inc. (PSTG) Q4 2024 Earnings Call Transcript*. Verfügbar 17. Mai 2024 unter <https://seekingalpha.com/article/4674588-pure-storage-inc-pstg-q4-2024-earnings-call-transcript>
- Gregory, R. P. (2021). What determines Manager and Investor Sentiment? *Journal of Behavioural and Experimental Finance*, 30, 100499.
- Gupta, R. (2019). Manager sentiment and stock market volatility. *Journal of Management Information and Decision Sciences*, 22(1), 11–21.
- Hales, J., Kuang, X. (, & Venkataraman, S. (2011). Who Believes the Hype? An Experimental Examination of How Language Affects Investor Judgments. *Journal of Accounting Research*, 49(1), 223–255.
- Hennig, J. C., Firk, S., & Wolff, M. (2023). Credibility Signals from Soft Information: Evidence from Investor Reactions to Tone in Earnings Conference Calls. *European Accounting Review*.
- Henry, E. (2008). Are investors influenced by how earnings press releases are written? *The Journal of Business Communication* (1973), 45(4), 363–407.
- Henry, E., & Leone, A. J. (2016). Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone. *The Accounting Review*, 91(1), 153–178.
- Hollander, S., Pronk, M., & Roelofsen, E. (2010). Does Silence Speak? An Empirical Analysis of Disclosure Choices During Conference Calls. *Journal of Accounting Research*, 48(3), 531–563.
- Holler, J. (2016). *Einführung in die Event Study Methodik*. Shaker Verlag.
- Huang, X., Teoh, S. H., & Zhang, Y. (2013). Tone management. *The Accounting Review*, 89(3), 1083–1113.
- Irani, A. J. (2004). The Effect of Regulation Fair Disclosure on the Relevance of Conference Calls to Financial Analysts. *Review of Quantitative Finance and Accounting*, 22(1), 15–28.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.
- Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, 132(1), 126–149.
- Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, 33, 171–185.
- Klevak, J., Livnat, J., & Soslava, K. (2024). Gender difference in CFO communication: Evidence from earnings calls. *Journal of Accounting and Public Policy*, 44, 107193.
- Lee, J. (2015). Can Investors Detect Managers' Lack of Spontaneity? Adherence to Predetermined Scripts during Earnings Conference Calls. *The Accounting Review*, 91(1), 229–250.
- Liu, B. (2020). *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions* (2. Aufl.). Cambridge University Press.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65.
- Loughran, T., & McDonald, B. (2016). Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research*, 54(4), 1187–1230.
- LSEG. (2024). LSEG ESG Scores. Verfügbar 17. August 2024 unter <https://www.lseg.com/en/data-analytics/sustainable-finance/esg-scores>
- Luo, R., & Ye, Y. (2023). Pressure from words: The tone of investors in Chinese earnings communication conferences and managerial myopia. *Accounting & Finance*, 64(1), 833–868.
- Luo, Y., & Zhou, L. (2017). Managerial ability, tone of earnings announcements, and market reaction. *Asian Review of Accounting*, 25(4), 454–471.
- Malkiel, B. G., & Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417.
- Mather, P., Ranasinghe, D., & Unda, L. A. (2021). Are gender diverse boards more cautious? The impact of board gender diversity on sentiment in earnings press releases. *Journal of Contemporary Accounting & Economics*, 17(3), 100278.
- Matsumoto, D., Pronk, M., & Roelofsen, E. (2011). What Makes Conference Calls Useful? The Information Content of Managers' Presentations and Analysts' Discussion Sessions. *The Accounting Review*, 86(4), 1383–1414.
- Mayew, W. J., Sethuraman, M., & Venkatachalam, M. (2020). Individual Analysts' Stock Recommendations, Earnings Forecasts, and the Informativeness of Conference Call Question and Answer Sessions. *The Accounting Review*, 95(6), 311–337.
- Milian, J. A., Smith, A. L., & Alfonso, E. (2017). An Investigation of Analysts' Praise of Management During Earnings Conference Calls. *Journal of Behavioral Finance*, 18(1), 65–77.
- Müller, S. (2023). Expected Return Models: Choosing the Right Approach. Verfügbar 24. Mai 2024 unter https://eventstudy.de/models/expected_return.html

- Ning, Y., Yang, S., & Xiao, Z. (2023). Manager sentiment, stock return, and the evolving information environment in post-IPO firms. *Journal of Economics and Finance*.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, 54(1), 547–577.
- Pouryousof, A., Nassirzadeh, F., & Askarany, D. (2023). Inconsistency in managers' disclosure tone: the signalling perspective. *Risks*, 11(12), 205.
- Price, S. M., Doran, J. S., Peterson, D. R., & Bliss, B. A. (2012). Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, 36(4), 992–1011.
- Rahman, S. (2023). Narrative tone and earnings persistence. *Journal of International Accounting, Auditing and Taxation*, 52, 100562.
- Renault, T. (2017). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking & Finance*, 84, 25–40.
- Rodriguez, A. (2019). Netflix plunges more than 10% after a huge miss on subscriber growth during Q2. Verfügbar 18. Juni 2024 unter <http://www.businessinsider.com/live-netflix-q2-2019-earnings-subscriber-growth-analysis-2019-7>
- Rogers, J. L., Van Buskirk, A., & Zechman, S. L. (2011). Disclosure Tone and Shareholder Litigation. *The Accounting Review*, 86(6), 2155–2183.
- S&P Global. (2024). *S&P Composite 1500®*. Verfügbar 17. Mai 2024 unter <https://www.spglobal.com/spdji/en/indices/equity/sp-composite-1500/>
- Stone, P. J., & Hunt, E. B. (1963). A Computer Approach to Content Analysis, Studies Using the General Inquirer System. *Managing Requirements Knowledge, International Workshop On*, 1, 241–256.
- Tama-Sweet, I. (2014). Changes in earnings announcement tone and insider sales. *Advances in Accounting*, 30(2), 276–282.
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139–1168.
- Todd, A., Bowden, J., & Moshfeghi, Y. (2024). Text-based sentiment analysis in finance: Synthesising the existing literature and exploring future directions. *International Journal of Intelligent Systems in Accounting, Finance & Management*, 31(1), 1–13.
- Tuffery, S. (2023). *Deep Learning: From Big Data to Artificial Intelligence with R*. John Wiley & Sons.
- Yamamoto, R., Kawadai, N., Kurita, M., & Baba, S. (2022). Managements' tone strategies by earnings call transcripts in the global markets. *Journal of Asset Management*, 23(3), 246–255.
- Yan, S., Hwang, I., Jung, H.-R., & Jung, T. (2023). Abnormal Disclosure Tone in MD&A and Analysts' Earnings Forecast: Evidence from China. *Emerging Markets Finance and Trade*, 1–17.
- Yang, B., Liu, Y., & Su, Y. (2022). Earnings communication conferences and post-earnings-announcement drift: Evidence from China. *Accounting & Finance*, 63(2), 2145–2185.
- Zhou, G. (2018). Measuring investor sentiment. *Annual Review of Financial Economics*, 10, 239–259.

Abkürzungsverzeichnis

BM Buchwert-zu-Markt

CAR(-1,1) Kumulierte abnormale Renditen im dreißigsten Fenster um den Earnings Call

CAR(-2,42) Kumulierte abnormale Renditen in der Driftperiode 2 bis 42. Tage nach dem Earnings Call

CEO Chief Executive Officer

CFO Chief Financial Officer

EMH Effizienzmarkthypothese

ESG Environmental, Social, and Governance (englisch für: Umwelt, Soziales und Unternehmensführung)

GI Harvard-GI-Lexikon

HE Henry-Lexikon

IO Anteil institutioneller Anleger

LM Loughran-and-McDonald-Lexikon

PEAD Post-Earnings Announcement Drift (englisch für: Drift nach der Gewinnankündigung)

Q&A Frage- und Antwortabschnitt eines Earnings Calls

Reg FD Regulation Fair Disclosure

ROA Return on Assets

Symbolverzeichnis

$\hat{\alpha}_i$ Geschätzte Konstante für die erwartete Rendite der Aktie i

$AnalystEPS_{i,q}$ Von Analysten erwarteter Gewinn pro Aktie des Unternehmens i im Quartal q

$ANALYSTS$ Anzahl am Earnings Call teilnehmender Analysten

$AnalystSURP_{i,q}$ Analysten-basierte Gewinnüberraschung des Unternehmens i im Quartal q

$AR_{i,t}$ Abnormale Rendite der Aktie i am Tag t

$\hat{\beta}_i$ Geschätzte Sensitivität der Aktie i gegenüber der Marktrendite $R_{m,t}$

BM Buchwert-zu-Markt-Verhältnis

$CAR(t_1, t_2)$ Kumulierte abnormale Renditen von t_1 bis t_2 nach dem Earnings Call des Unternehmens i im Quartal q	$Sentiment_{Gesamt(i,q)}$ Sentiment in dem gesamten Earnings Call
$\sum Controls_{i,q}$ Vektor von Kontrollvariablen auf Unternehmensebene im Quartal q	$Sentiment_{Manager(Q\&A,i,q)}$ Sentiment der Manager in dem Q&A
δ_i Feste Effekte für das Unternehmen i	$Sentiment_{Präsentation(i,q)}$ Sentiment in der Präsentation
$EPS_{i,q}$ Gewinn je Aktie des Unternehmens i im Quartal q	$Sentiment_{Q\&A(i,q)}$ Sentiment in dem Q&A
$\varepsilon_{i,t}$ Störterm der Regression für Unternehmen i im Quartal q	$SIZE$ Logarithmierte Marktkapitalisierung in Millionen US-Dollar
$EPS_{i,q-4}$ Gewinn je Aktie des Unternehmens i im gleichen Quartal des Vorjahres	$SURP_{i,q}$ Gewinnüberraschung des Unternehmens i im Quartal q
$E(R_{i,t})$ Erwartete Rendite der Aktie i am Tag t	t Zeit
$FIRMAGE$ Jahre seit dem Börsengang	
$GROWTH$ Prozentuales Umsatzwachstum	
i Unternehmen	
l Lexikon (LM; Henry; Harvard-GI)	
λ_q Feste Effekte für das Quartal q	
$LEVERAGE$ Verschuldungsgrad	
$LOSS$ Dummy-Variable für Verluste	
$N_{negativ}$ Anzahl negativ klassifizierter Wörter	
$N_{positiv}$ Anzahl positiv klassifizierter Wörter	
q Quartal	
$R_{i,t}$ Tatsächliche Rendite der Aktie i am Tag t	
$R_{m,t}$ Rendite des Marktportfolios am Tag t	
ROA Return on Assets	
$\bar{R}_{p,t}$ Mittlere Rendite eines Portfolios von Aktien mit dem gleichen Größen- und Buchwert-zu-Markt-Quartil wie Unternehmen i	
$Schlusskurs_{i,t}$ Schlusskurs der Aktie i am Tag t	
$Schlusskurs_{i,t-1}$ Schlusskurs der Aktie am Vortag	
$Schlusskurs_{i,q-4}$ Schlusskurs der Aktie des Unternehmens i am Ende des Quartals des Vorjahres	
$Sentiment_{i,q}$ Sentiment	
$Sentiment_{Analyst(Q\&A,i,q)}$ Sentiment der Analysten in dem Q&A	