The Imitation Game

A Longitudinal Study of the Imitation-Performance Relationship in S&P 500 Firms from 2009-2023

William Ericson
Erik Lang

Supervision:
Sebastian Krakowski

Thesis for the Degree of Master of Science
MSc. in Business & Management
Stockholm School of Economics, Spring 2024

Abstract

We analyse the imitation behaviour of S&P 500 companies and its link to diverse financial performance outcomes during the period 2009-2023. We measure imitation with multiple variables based on communication indicators derived from the annual 10-K disclosures of all companies observed in the study. These variables are then used in a series of panel regressions that establish statistically significant relationships between imitation behaviour and multiple dimensions of firm performance, namely stock returns, revenue growth and profitability growth. We overcome the lack of robust and diverse empirical evidence and show that imitation behaviour positively affects firm performance. A positive imitation-performance relationship moderates tensions between the RBV and dynamic capability literature, and our results support viewing imitation as a key enabler of dynamic capabilities. As a by-product of the methodological choices of machine-learning-enabled sentiment analysis and LDA (Latent Dirichlet Allocation) topic modelling on 10-K annual reports, the study also contributes with the LDA-created topic dictionaries. Each dictionary for each year from 2009 to 2022 contains the 100 statistically most important words S&P 500 companies use when speaking about innovation, uncertainty, and sustainability.

Keywords: Imitation, Firm performance, Dynamic capabilities, LDA topic modelling, ML sentiment analysis, S&P 500, 10-K annual report

Acknowledgements

We express our sincere gratitude to Professor Sebastian Krakowski for his supervision throughout the process of this master's thesis. We would also like to extend our sincere appreciation to Professor Alexander Mafael for his methodological advice, which significantly enriched the quality of our study. Sebastian's and Alexander's contributions have been invaluable, and we are truly thankful for their dedication and support: Both for academic support on this thesis and personal guidance on the careers we are aiming for.

We furthermore want to extend our appreciation to Professor Susanne Sweet and Professor Karl Wennberg, for organising the thesis seminar and providing us with the fundamental knowledge to execute this study.

Lastly, we are grateful for the Stockholm School of Economics' academic training, learning environment and logistical support, having made both this thesis and the last two years a highly enjoyable, intellectually stimulating and academically rewarding experience.

Stockholm, 2024

William Ericson and Erik Lang

Abbreviations

CIK: Central Index Key
CRSP: Center for Research in Security Prices, LLC
CSR: Corporate Social Responsibility
DIS: Disruptive Innovation Score
DW: Durbin-Watson
EBIT: Earnings before interest and taxes
EDGAR: Electronic Data Gathering, Analysis, and Retrieval
E.-m. mentions: Early-mover mentions
E.-m. sen. breadth: Early-mover sentiment breadth
E.-m. sentiment: Early-mover sentiment
ESG: Environmental, Social, and Governance
IQR: Interquartile Range
Im. mentions: Imitation mentions
Im. sen. breadth: Imitation sentiment breadth
Im. sentiment: Imitation sentiment
LDA: Latent Dirichlet Allocation
MLR: Multiple Linear Regression
OLS: Ordinary Least Squares
Q-Q plot: Quantile-Quantile plot
RBV: Resource-based View
ROA: Return on Assets
S&P: Standard & Poor's
SEC: U.S. Securities and Exchange Commission
R^2: Coefficient of determination
VADER: Valence Aware Dictionary for sEntiment Reasoning
VIF: Variance Inflation Factors
WRDS: Wharton Research Data Services

Contents

Abstract ii
Acknowledgements iii
Abbreviations iv

1 Introduction **1**
1.1 Context 1
1.2 Contributions 2
1.3 Delimitation 4
1.4 Structure 6

2 Conceptual Background **7**
2.1 Strategic Imitation 7
2.2 Adjacent Constructs 10
2.3 Conceptual Framework 15

3 Empirical Study **20**
3.1 Research Design 20
3.2 Data & Context 23
3.3 Variables 27
3.4 Descriptive Statistics 37
3.5 Model Specification 39
3.6 Results 45
3.7 Quality Discussion 49

4 Analysis & Discussion **53**
4.1 Imitation-Performance Relationship 53
4.2 Sentiment Deviation 56
4.3 Post-Hoc 57

5 Conclusions **61**
5.1 Theoretical Contributions 62
5.2 Managerial Implications 64

5.3 Limitations 65
5.4 Future Research 66

Bibliography **68**

Appendix **77**
OLS Regression Types 77
LDA Dictionaries 77
Note on Generative AI 97

1

Introduction

This chapter contains (1) the contextualisation of the research topic, (2) the research objective of the study and its contributions, (3) the delimitation of the study and theoretical scope and (4) the research structure and approach.

1.1 Context

Human beings, by their very nature, are social creatures, and this inherent sociality has played a crucial role in our evolutionary history. From the earliest days of hunter-gatherer societies to the complex world of the global economy, the tendency to align our actions and decisions with those of others – a phenomenon known as herding – has been a constant presence. This instinct, deeply embedded in our psyche, drives us to look to our peers for cues on how to behave, especially in situations of uncertainty or ambiguity. Herding behaviour has been discussed as far in the past as Keynes (1936) and extensively explored in economics and behavioural science, leading to inefficient decisions (Banerjee, 1992). Simultaneously, a closely related theoretical construct has been a central pillar within strategy research, the resource-based view, and the dynamic capability literature: Imitation. Imitation is the replication of another firm's products, processes, capabilities, technologies, or decisions and the prevalent static explanation of why firm heterogeneity and competitive advantages erode (Barney, 1991; Peteraf, 1993). Contradictory scholars recently proposed that imitation may also be a key driver of innovation and result in a sustained competitive advantage (Posen et al., 2023). This counterintuitive suggestion is rooted in the re-evaluation of the underlying assumptions within the imitation literature and aims to move from a static imitation model with weakening effects on rare resources to an evolutionary and dynamic model where imitation becomes a central enabler of dynamic capabilities (Posen et al., 2023). While Posen et al. (2023) explicitly do not infer superiority from one model over the other, they highlight that a robust understanding

of the imitation-performance relationship is the substantial enabler to advance strategy research's understanding of imitation and ground future research to resolve the overarching theoretical frictions between the static and dynamic model including the underlying assumptions. However, despite the central role of the imitation-performance relationship, Posen, Lee, et al.'s (2023) literature review of 338 publications in 15 well-reputed management journals reveals that only a fraction of papers tried to answer this fundamental performance question and seldom test the imitation–performance relationship empirically, while even concluding in contradicting findings (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022). Consequently, a better understanding of the imitation-performance relationship becomes essential to inform strategy theorising: This is both relevant to intentional imitation in the pursuit of dynamic capabilities, knowledge recombination and innovation (Posen et al., 2023), but also related management fields that raise concerns about the performance effects of unintended herding and imitation (Nauhaus et al., 2021). This white space of understanding of the imitation-performance relationship furthermore extends into the finance literature, where scholars theoretically tend to agree on the causes of investor herding and imitation (Komalasari et al., 2022; Ramli et al., 2015; Zhou and Lai, 2009), yet fail to conclusively assess whether it leads to better or worse investment decisions (BenMabrouk, 2018; Chen and Demirer, 2018; Demirer and Zhang, 2019; Komalasari et al., 2022; Singh, 2013).

This cross-disciplinary empirical uncertainty and theoretical ambiguity between the RBV and dynamic capability literature on the expected performance outcomes of imitation calls for an answer to the question: Does imitation behaviour affect firm performance?

1.2 Contributions

In this study, we answer the research question: Did imitation behaviour of S&P500 companies in innovation contexts affect firm performance during the period 2009-2023?

The following subsection describes the contributions of how (1) imitation positively affects diverse performance measures, overcoming frictions between the static and dynamic model of imitation, and fostering a dialogue between management and economics research. (2) We overcome the lack of robust and diverse empirical evidence on the performance outcomes of imitation. (3) We methodologically contribute to the management field by developing partly novel imitation measures and employing machine-learning-enabled sentiment analysis combined with Latent Dirichlet Allocation (LDA) topic modelling on 10-K annual reports. (4) We provide the management scholars with the LDA-created topic dictionaries, covering the 100 most important words companies use when speaking

about innovation, uncertainty, and sustainability for each year from 2009-2022. Lastly, (5) we demonstrate post-hoc how being an early-mover in a sustainability context yields diverse performance improvements.

First, our results show that a higher degree of imitation behaviour has a positive effect on firm performance. Strategic imitation is the "attempt to reproduce [...] another firm's product, processes, capabilities, technologies, or [strategic] decision [...]" (Posen et al., 2023, p 76). The positive imitation-performance relationship can be explained by applying a dynamic capability perspective (Teece et al., 1997), viewing imitation not merely as a reactive strategy but as an enabler of dynamic capabilities itself (Posen et al., 2023) serving as a mechanism for firms to acquire and integrate external innovations that complement their existing capabilities, thereby fostering an environment of continuous learning and adaptation on an aggregated market level (Cohen and Levinthal, 1990; Levitt and March, 1988). A positive imitation-performance relationship furthermore aligns with Posen et al.'s (2023) dynamic model of imitation, where the imitator integrates its own and competitors' knowledge, processes, and technologies and thus avoids potential adverse effects of disregarding private information as described by economic scholars (Banerjee, 1992; Bikhchandani et al., 1992). The positive imitation-performance relationship seeks to be a starting point for a variety of research questions within imitation research (Posen et al., 2023).

Second, we empirically overcame the lack of robust and diverse evidence on the performance outcomes of imitation and bridged tensions between the static and dynamic imitation models hindering effective measurement (Miller and McKee, 2021; Posen et al., 2023; Samba et al., 2021). We showed that approaching the imitation-performance relationship from a firm-centric aggregated performance perspective, measuring imitation as the observable correlated behaviour of a group of organisations, combined with statistically robust panel regression models (Berry, 1993; Torres-Reyna, 2007; Wolf and Best, 2013) foster academic discourse. We filled the research gap by accounting for diverse performance outcomes, including shareholder value creation, revenue growth, and profitability growth for S&P 500 companies from 2009-2022 (Posen et al., 2023).

Third, we methodologically contribute to management research by developing partly novel communication indicators to measure imitation behaviour while utilising large-scale machine-learning-enabled LDA and sentiment analysis to extract information value from textual data (Bayer et al., 2017; Bellstam et al., 2021; Blei et al., 2003; Li et al., 2013; Ligorio et al., 2022; Miric et al., 2023). Our study furthermore serves as an example that pre-existing dictionaries potentially cause issues when transferred to other documents or time periods (Miric et al., 2023). Compared to the LDA-created innovation dictionary, the Disruptive Innovation Score (DIS) dictionary, containing 500 pre-defined keywords

around 29 disruptive technologies (Nguyen and Agrawal, 2023), struggled to produce significant results.

Fourth, we provide researchers with our dictionaries created via LDA topic modelling. The dictionaries in Appendix contain, for each year from 2009-2022, the 100 most important words S&P 500 companies use when speaking about innovation, uncertainty, and sustainability.

Fifth, we post-hoc show how our early-mover variable in a sustainability context contributes to diverse performance improvements. Firms that address sustainability in negatively connotated contexts also experience positive performance outcomes, according to our early-mover sentiment variable.

In summary, we contribute to the academic discussion by answering the research question and providing evidence that imitation behaviour positively affects diverse performance outcomes. While centred around the theoretical questions of the imitation-performance relationship, we believe the created dictionaries can contribute to various management fields.

1.3 Delimitation

In this study, we focus on the imitation-performance relationship and the fundamental issue of whether imitation is a driving force of performance as proposed by the dynamic capability view or rather an erodent of competitive advantage and antecedent of improved firm performance as expected by the Resource-based View (RBV). When doing so, we (1) do not address other fundamental issues in strategic imitation research, (2) we take an aggregated outcome lens and not a dyadic perspective with (3) objective and diverse financial performance measures in (4) a distinct innovation context.

First, we do not address other fundamental issues in strategy research related to imitation (Posen et al., 2023) such as the dimensions and characterisations of uncertainty and its impact on imitation (Rumelt, 1984). While we post-hoc extend the context of our studies from innovation into sustainability and uncertainty via our additional LDA-created dictionaries, we do not include uncertainty or sustainability measures. The measurement of imitation remains unaffected by the context. Next, we do not investigate the industry conditions under which the contradicting assumptions between the static and dynamic assumption of imitation hold (Sharapov and Ross, 2023), and we do not investigate imitation in combination with other fundamental strategy rules (Eisenhardt and Martin, 2000; Eisenhardt and Sull, 2001), as we only measure if a firm imitates or not and to what degree. Furthermore, we do not explore the linkage between imitation strategies and intra-organisational behaviour (Knott and Turner, 2019), as we focus on the imitation

outcome and not the process of how firms internally derive their strategic decision if, how and how much they want to imitate their competitors and the market. We do not explore the link between imitation and concrete crisis times, such as specific responses related to, e.g., the COVID-19 pandemic (Lee et al., 2020), while only starting our observation period after the financial crisis of 2008. Lastly, we do not research the role of imitation for entrepreneurship (McDonald and Eisenhardt, 2020; Young and Barney, 2018) as we focus on existing large corporations spanning more than 80% of the U.S. market capitalisation (see Section 3.2.1).

Second, when exploring the imitation-performance relationship, we do not take a dyadic lens but investigate the imitation-performance relationship from an aggregated performance perspective. Consequently, we also do not investigate leader-follower asymmetries (Haveman, 1993) or to what extent specific capabilities enable imitation strategies (Cohen and Levinthal, 1990). Furthermore, we do not focus on the role of imitation in concrete global business challenges such as foreign direct investments (Henisz and Delios, 2001; Yu and Ito, 1988), the timing of expansion (Delios et al., 2008) or market entry decisions (Guillén, 2003).

Third, when investigating the imitation-performance relationship, we focus on objective and diverse financial performance outcomes. We do not explore subjective performance outcomes or Environmental, Social, and Governance (ESG) and Corporate Social Responsability (CSR) outcomes. We do this as subjective performance outcomes tend to hinder progress when investigating fundamental strategy issues (Miller and McKee, 2021; Samba et al., 2021), such as imitation, being at the core of the RBV and dynamic capability literature. We do not include ESG and CSR outcome metrics as rating methodologies change over time and might not capture the true ESG and CSR impact of the firm (Escrig-Olmedo et al., 2019) resulting in inconsistent outcome measures for our longitudinal observation period from 2009-2023. Furthermore, unfortunately our accessible Wharton Research Data Services (WRDS)-subscribed data sources do not include sufficient ESG and CSR databases to generally support a robust inclusion of ESG and CSR performance measures.

Fourth, we study the imitation-performance relationship in an innovation context as innovation is theoretically at the core of the imitation literature (Barney, 1991; Peteraf, 1993) and studied throughout the strategy literature (Posen et al., 2023). Innovation is characterised by the utilisation of technologies, products, services, business processes, or business models that improve performance (Christensen, 1997; Nguyen and Agrawal, 2023; Zhou, Yim, et al., 2005). While the focus of our study centres around the imitation-performance relationship in the innovation context, our methodological choice of employing LDA topic modelling, a type of statistical machine learning model used in natural language processing and text mining (Blei et al., 2003), allows us post-hoc to extend our research

into the contexts of uncertainty and sustainability. We highlight that these contexts only extend our imitation measures and are, e.g. not a time or event-bound quantification of the amount of uncertainty in an organisation's environment.

In summary, our thesis focuses on the fundamental relationship between imitation and performance, specifically examining whether imitation drives or undermines performance. We intentionally exclude other dimensions of strategic imitation, such as detailed internal processes and broader socio-economic impacts, adopting a macro-level view on aggregated financial outcomes within an innovation context.

1.4 Structure

This thesis is organised into five sections that collectively explore the imitation-performance relationship within the strategic management literature. Each section is structured to progressively build upon the insights developed in the preceding sections, synthesising theoretical constructs and empirical evidence to offer a comprehensive analysis of the imitation-performance relationship.

Within the main five sections, we describe how we overcome the combined nascent friction within the static and dynamic understanding of imitation and a lack of empirical evidence on the imitation-performance relationship (Posen et al., 2023) by developing hypotheses based on the RBV, dynamic capability, and economic herding literature (Banerjee, 1992; Barney, 1991; Bikhchandani et al., 1992; Peteraf, 1993; Teece et al., 1997) to empirically explore (Bell et al., 2022; Edmondson and Mcmanus, 2007) the imitation-performance relationship. As previous empirical evidence is both scarce and contradicting (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022; Posen et al., 2023), we utilise large-scale machine-learning enabled LDA and sentiment-based methods (Bayer et al., 2017; Bellstam et al., 2021; Blei et al., 2003; Li et al., 2013; Ligorio et al., 2022; Miric et al., 2023) in combination with objective and diverse performance measures such as shareholder value creation, revenue growth and profitability improvements (Berk and DeMarzo, 2007; Eberhart et al., 2004) in statistically robust pooled, fixed-effects and random-effects panel regressions (Berry, 1993; Torres-Reyna, 2007; Wolf and Best, 2013) to overcome frictions between established management theories (Miller and McKee, 2021; Samba et al., 2021) and estimate the positive imitation-performance and negative early-mover-performance relationship in innovation context. After analysing and discussing our main results, we demonstrate post-hoc how, in contradiction to innovation contexts, companies can be financially rewarded for addressing sustainability topics.

2

Conceptual Background

This chapter contains (1) an overview of imitation through the lens of the RBV and dynamic capabilities literature revealing the theoretical frictions on the expected imitation-performance outcomes, (2) adjacent theoretical constructs in the strategy literature as well as herding theories and how they support an aggregated performance perspective and offer new theoretical arguments to assess the imitation-performance relationship, and (3) the theorising of our conceptual framework including our hypotheses to resolve the frictions within the discussed theories and better understand the imitation-performance relationship.

2.1 Strategic Imitation

Imitation is a central pillar in strategy research and has been defined as the "attempt to reproduce [...] another firm's product, processes, capabilities, technologies or [strategic] decision [..]" (Posen et al., 2023, p. 76). The concept itself emerged as the erodent of a firm's competitive advantage (Barney, 1991; Lippman and Rumelt, 1982; Peteraf, 1993) and is found at the core of a variety of foundational strategy theories (Posen et al., 2023) such as the RBV (Barney, 1991) and dynamic capability literature (Teece et al., 1997). Furthermore, in the strategy literature, imitation is theoretically adjacent to the concepts of absorptive capacity (Cohen and Levinthal, 1990), diffusion (Strang and Soule, 1998), knowledge spill-overs (Jaffe, Trajtenberg, and Fogarty, 2000), mimetic isomorphism (DiMaggio, Powell, et al., 1983) and vicarious learning (Levitt and March, 1988). We furthermore argue that imitation is closely related to economic and financial herding theories (Banerjee, 1992; Komalasari et al., 2022) and that those existing theories and adjacent literature streams offer novel insights on the imitation-performance relationship. These insights can resolve frictions within the static and dynamic imitation literature and consequently advance management scholars' theoretical understanding of the imitation-performance

relationship while suggesting an aggregated performance lens that aligns with both the static and dynamic models of imitation.

The RBV posits that the basis for a sustained competitive advantage lies in the application of a bundle of resources – tangible and intangible – that enable a firm to perform better than its competitors (Barney, 1991). According to Barney (1991), resources must be valuable, rare, inimitable, and non-substitutable. Dynamic capabilities are strategic and organisational processes that allow firms to adapt to rapidly changing environments, thereby achieving and sustaining competitive advantage (Fleming et al., 1996). Dynamic capabilities extend the RBV by focusing not just on the resources a firm possesses but also on its ability to reshape and reconfigure these resources in response to changing external demand, vital for organisations operating in industries characterised by rapid technological change (Eisenhardt and Martin, 2000). The core components of dynamic capabilities are the ability to sense opportunities and threats, the capacity to seize opportunities, and the ability to transform available resources to maintain competitiveness (Teece et al., 1997).

When examining the concept of imitation in innovation and technological change contexts (e.g. Davis and Aggarwal, 2020; Giachetti, Lampel, et al., 2017), it becomes evident that the RBV and dynamic capability literature expect dichotomous outcomes (Posen et al., 2023). From an RBV perspective, and amidst the proposition of inimitability, imitation reduces the rareness and value of resources, and subsequently firm heterogeneity and competitive advantage, for both the imitator and the imitation-target (Barney, 1991). On the other hand, and from a dynamic capability perspective, imitation might be an enabler of competitive advantage when the imitator reproduces some resources of the target and either on purpose or by accident combines these with some of its own to create innovation, valuable resources and promote firm performance (Posen, Lee, et al., 2013; Posen and Martignoni, 2018; Teece et al., 1997; Wang et al., 2023).

Posen et al.'s (2023) literature review of 338 publications in 15 well-reputed management journals synthesises these tensions in a static and dynamic understanding of imitation by contrasting the differences in the underlying assumptions of imitation behaviour. The static model assumes that imitation is easy, weak firms imitate, and uncertainty in the firm's environment fosters imitation (Posen et al., 2023). The static model furthermore understands imitation as the singular knowledge transfer from the target to the imitator (Posen et al., 2023). On the other hand, the dynamic model implies that imitation is difficult, strong firms imitate, uncertainty hinders imitation and that many imitation strategies exist (Posen et al., 2023). The dynamic model furthermore describes that the imitator recombines the target knowledge, practices and decision with one's own while being simultaneously targeted by other imitators (Posen et al., 2023). The dynamic model thus not only captures the dyadic relationship between two actors but a market-level perspective where change is constant, and imitation is an ever-ongoing process (Posen

et al., 2023). In the dynamic model, individual firms make strategic decisions to which extent they aim to imitate and recombine knowledge to create innovation (Posen et al., 2023). Consequently, and contradictory to the RBV, the dynamic model challenges the predominant view in strategy research that imitation reduces inter-firm performance heterogeneity (Posen et al., 2023). The dynamic model predicts increased performance heterogeneity and argues that imitation is a key driver of dynamic capabilities, innovation and competitive advantage (Posen et al., 2023). While Posen et al. (2023) explicitly do not infer superiority from one model over the other, they highlight that a robust understanding of the imitation-performance relationship is the substantial enabler to advance the strategy research's understanding of imitation, its underlying assumptions and ground future research to resolve the overarching theoretical frictions between the static and dynamic model (Posen et al., 2023).

Posen et al.'s (2023) key literature review not only highlights this theoretical friction but also calls for an answer to the question of whether imitation increases or decreases performance outcomes. Strategy literature is lacking a substantial understanding of this relationship (Posen et al., 2023), and seldomly test the imitation–performance relationship empirically (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022). The existing and contradicting evidence on the imitation-performance relationship furthermore relies on small-sample studies in single industries dating back multiple decades: An observation of 30 telecommunication firms from 1996 to 2005 showed that imitation and investments in similar technologies are beneficial to all firms in the industry (Banker et al., 2013). A similar observation was made from 1997-2008 when UK mobile phone vendors experienced better sales performance when imitating their peers (Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022). On the other hand, imitation among Portuguese bank branches between 1988 and 1996 resulted in negative profitability effects (Barreto and Baden-Fuller, 2006). Furthermore, strategic similarity among US banks during the observation period from 1985-1992 resulted in negative performance effects (Deephouse, 1999). Consequently, the lack of recent, large-scale and robust empirical evidence for the dis-equilibrating power and performance improvements of imitation proposed by the dynamic model is a significant gap in strategy research. Filling this gap would provide fundamental guidance on the nature of the imitation-performance relationship. A positive imitation-performance relationship would consequently challenge the predominant view on imitation in large parts of strategy research (Posen et al., 2023). It would further nudge the literature to relax the implied assumption within the static understanding and rethink the theoretical foundation of imitation (Posen et al., 2023).

While the dynamic capability-based understanding of imitation challenges the predominant view on the imitation-performance relationship, we argue that various con-

structs adjacent to strategic imitation suggest an aggregated performance lens, illuminate its underlying mechanics, and consequently expand our theoretical understanding of how firms behave to robustly assess the imitation-performance relationship.

2.2 Adjacent Constructs

Posen et al.'s (2023) literature review of 338 publications in 15 well-reputed management journals links imitation to a variety of adjacent theoretical constructs (Cohen and Levinthal, 1990; DiMaggio, Powell, et al., 1983; Jaffe, Trajtenberg, and Fogarty, 2000; Levitt and March, 1988; Strang and Soule, 1998). By doing so, they acknowledge that imitation is not defined within a bounded literature and that "relevant contributions may be hidden under different but related labels" (Posen et al., 2023, p 77). We therefore propose that imitation is also closely related to herding theories as widely spread in the economics and finance literature (Banerjee, 1992; Komalasari et al., 2022). By considering constructs adjacent to strategic imitation, we argue for a more comprehensive theoretical understanding of the imitation-performance relationship. This better understanding is enabled by an aggregated performance lens and new perspectives on the expected outcomes of the imitation-performance relationships offered by economic herding theories. The adjacent theoretical constructs provide deeper insights into how firms try to leverage external knowledge and innovations from the aggregated market while herding theories offer insights to resolve the frictions between the static and dynamic understanding of the imitation-performance relationship.

Therefore we introduce the reader to (1) the reasoning behind our aggregated performance lens by highlighting links between strategic imitation and a variety of adjacent constructs, (2) central herding theories and how they offer potential insights to resolve the imitation-performance frictions to provide management scholars with novel theoretical insights, and (3) and introduction to the early-mover discussion, as the strategic opposite to imitation.

First, our investigation of imitation on an aggregated level allows us to capture the outcomes of both the static and dynamic imitation models without being limited by the underlying theoretical frictions of the different models. While the static model originated from a dyadic understanding of imitation between two actors, Posen et al.'s (2023) large-scale literature review revealed that most of the static-model theorists have not distinguished between the dyadic target and imitator. Thus, we also do not track dyadic relationships but focus on how much each firm imitates (or not). Consequently, in a theoretical static understanding, we no longer need to evaluate the outcomes of each individual static relationship but explore the overall aggregate outcome of all summed static imitations over a given time period. By doing so, we theoretically allow for the static

assumptions to hold and only analyse whether, on aggregate, an individual company experiences positive or negative outcomes. The dynamic model furthermore views imitation as an ongoing market-level process where all actors imitate (and are being imitated by) multiple actors at the same time: Change is the only constant, and imitation is an ever-ongoing process (Posen et al., 2023). In consequence, measuring the extent to which each company engages in imitation and whether they experience positive or negative performance is again aligned with the aggregated lens. Using an aggregated lens to explore the imitation-performance relationship is furthermore supported by the adjacent constructs of absorptive capacity (Cohen and Levinthal, 1990), diffusion (Strang and Soule, 1998), knowledge spill-overs (Jaffe, Trajtenberg, and Fogarty, 2000), vicarious learning (Levitt and March, 1988) which are accounting for dynamic macro-processes between multiple actors and market participants.

Absorptive capacity is defined as "the ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends," which is critical to its innovative capabilities (Cohen and Levinthal, 1990, p 128). It is intrinsic to a firm's ability to imitate strategically. Absorptive capacity acts as a facilitator of imitation (Davis and Aggarwal, 2020). Thus, absorptive capacity takes a firm-centric macro-level view, rather than taking the dyadic perspective, and implicitly assumes the existence of external knowledge largely available as a source for the entire market (Cohen and Levinthal, 1990).

Diffusion pertains to the spread of elements like behaviours, strategies, technologies, or structures within a social system (Strang and Soule, 1998). This focus aligns with strategic imitation as both processes involve the uptake and implementation of external innovations within a firm (Posen et al., 2023). In the context of strategic imitation, diffusion represents a potential market-level outcome. Thus, diffusion again assumes an aggregated outcome lens, as behaviours, strategies, technologies, or structures spread between various actors.

Knowledge spillovers occur when investments in knowledge creation by one actor facilitate innovation by others without full compensation, effectively creating a pool of accessible knowledge (Jaffe, Trajtenberg, and Fogarty, 2000). Spillovers can be a result of or a facilitator for strategic imitation (Posen et al., 2023). Similar to diffusion, knowledge spillovers affect the aggregate market and not just the dyadic actors.

Mimetic isomorphism describes how organisations model themselves after peers they perceive as more legitimate or successful (DiMaggio, Powell, et al., 1983). This construct provides a motivation for strategic imitation: seeking legitimacy. The ubiquity of certain organisational structures or strategies, often adopted not because they enhance efficiency but because they are viewed as legitimate, highlights the normative pressures that drive

imitation (Posen et al., 2023). Again, the process of legitimacy-seeking is not limited to a single actor within the market.

Vicarious learning involves learning from the experiences of others without bearing the full costs of exploration and experimentation (Levitt and March, 1988; Terlaak and Gong, 2008). While vicarious learning includes imitation, it also encompasses broader learning activities that may not involve direct imitation. Firms often use the insights gained from others' experiences to inform their own strategic decisions, which might include but are not limited to imitation. The relationship between vicarious learning and strategic imitation illustrates the spectrum of learning mechanisms that firms employ to enhance their competitiveness (Posen et al., 2023). Once more, vicarious learning implies a dynamic process between multiple actors within the entire market.

Second, while the word "Herding" is not mentioned in Posen et al.'s (2023) literature review and furthermore not a common concept in management literature in general[1], we find direct references to central economic herding papers (Banerjee, 1992; Bikhchandani et al., 1992) and numerous implicit theoretical overlaps within the assumptions of Posen et al.'s (2023) static and dynamic model of imitation. By applying economic herding theories to the strategic imitation-performance literature, we reveal contrasting expectations about the performance outcome of the static imitation model but find support for the expected positive outcome of the dynamic model.

Herding, in its most general form, is defined as "behaviour patterns that are correlated across individuals" (Devenow and Welch, 1996). It is widely discussed in behavioural sciences and refers to a form of convergent social behaviour and alignment of individuals in a group without centralised coordination (Raafat et al., 2009). In economic contexts, this phenomenon has already been discussed as far in the past as Keynes (1936). A fundamental principle of classical economic theory is that investment choices are based on rational expectations utilising all accessible information, yet agents do not always behave in such a rational way and sometimes disregard their own information and mimic the behaviour of their peers (Scharfstein and Stein, 1990). Banerjee (1992) was one of the first to explain how such herding leads to inefficient decisions. According to his sequential decision model, each decision-maker considers the decisions made by those before them when making their own decision. This is rational because previous decision-makers may possess information that could be pertinent. However, the decision rules adopted by individuals are characterised by herd behaviour, meaning people tend to do what others are doing rather than using their own information. As a result, the equilibrium achieved through this herding behaviour is inefficient because it disregards the private information of individuals, which could lead to better decision-making outcomes if they were utilised (Banerjee, 1992).

[1]A normal search for herding within management or strategy literature yields little if no results.

This explanation of herding is also referred to as an information cascade (Bikhchandani et al., 1992). Recently, herding has furthermore become a popular concept in the finance literature (Komalasari et al., 2022) where investor herding is defined as the observable correlated behaviour of a group of investors, occurring consciously or unconsciously as they trade in the same direction at the same time (Hwang and Salmon, 2004; Nofsinger and Sias, 1999).[2] Similar to management research, finance scholars theoretically are not conclusive in the outcome assessment, yet in significant cases, investor herding also has a negative connotation due to the disregard of private information (BenMabrouk, 2018; Chen and Demirer, 2018; Demirer and Zhang, 2019; Komalasari et al., 2022; Singh, 2013). For example, fund managers who followed the market to avoid reputational risk underperform their non-herding peers over various time horizons (Jiang and Verardo, 2013).

Utilising the enhanced understanding of economic herding, we can reevaluate the strategic imitation-performance relationship and provide theoretical arguments as to why economic scholars would reject the static imitation model and support the expected positive performance outcomes of the dynamic model.

For the static model of imitation, and amidst the overall negative performance implications according to the RBV, all imitators are expected to converge to the global performance maximum (Levinthal, 1997, 2011; Rivkin, 2000). This assumption is a direct contradiction to economic herding theories and information cascade outcome, where herders, or here imitators, would settle on a suboptimal equilibrium by imitating and disregarding their own information (Banerjee, 1992; Bikhchandani et al., 1992). On the other hand, scholars who study dynamic imitation, while generally predicting positive imitation effects, also acknowledge the negative implications of information cascades. Similar to herding theories, they describe how ignoring seemingly inferior knowledge currently undervalued by the aggregated market may be the main cause of incumbent failure in times of disruptive innovation (Christensen and Bower, 1996). When imitations scholars argue that firms fail to realise performance benefits by not imitating strategically weaker or less dominant competitors, they support the viewpoint of herding theorists, where firms settle for a suboptimal equilibrium by not considering all available information and remaining with the current status quo and thus missing out on better outcomes (Banerjee, 1992; Bikhchandani et al., 1992). According to economic herding theories, strategic imitators are expected to achieve better outcomes because they combine both private and public information rather than disregarding their private information. In

[2] Here we do not refer to investor herding as a short-term process of fluctuating asset prices, where investors buy the same asset and temporarily drive prices up in a bubble-like manner. Instead, we mean when investors pursue similar long-term investment opportunities, which splits the real profits generated by an asset (or e.g. industry) among more investors, resulting in worse performance outcomes for all participating investors.

management theories, this is a direct paraphrasing of the dynamic imitation model, where the imitator integrates its own knowledge, processes, and technologies with those of multiple competitors.

We furthermore support our application of economic herding theories based on the general overlaps between the two schools of thought.

For example, the assumption of the static model, that imitation is easy and becomes normalised as more social actors adopt it (Haveman, 1993) and that imitation is easy when competitors' behaviours are observable, leading to many firms making similar strategic decisions (Haunschild and Miner, 1997; Keister, 2004; Simon and Lieberman, 2010) aligns with economic herding, where observability as synonym for ease is also an implied assumption. Moreover, uncertainty is expected to promote both imitation and herding behaviour (Bikhchandani et al., 1992; Posen et al., 2023). The assumption that weak firms often imitate without strategic intent, believing they will benefit from others' practices, particularly in uncertain situations (DiMaggio, Powell, et al., 1983; Vasudeva, 2013) as an automatic, mindless process resembles herd behaviour (Korn and Baum, 1999). Another example is the economic explanations for herding through principal-agent problems (Devenow and Welch, 1996). Principal-agent scenarios, where agents act on behalf of principals with potentially different incentives, can lead to herding behaviour as managers seek to protect their reputation or demonstrate competence (Devenow and Welch, 1996; Ross, 1973). This legitimacy-seeking motive is also evident in the imitation literature, where firms may imitate to provide legitimacy for potential mistakes, leading to worsened performance outcomes (Dye et al., 2014). Lastly, strategic imitation scholars tend to refer to financial herding concepts when strategic decisions are interdependent with competitors' actions or when imitation shapes rivals' decisions, mirroring investor behaviour in financial markets (Argyres et al., 2015; Sharapov and Ross, 2023). In summary, imitation perspectives often align with herding theories and partially go as far as using imitation as a synonym for herding, which supports our conclusion that herding theories help to resolve the theoretical frictions between the static and dynamic models of imitation.

Third, early-mover effects describe the strategic advantages and costs of being first or among the first when e.g. entering a market or adopting a new technology (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). From the RBV, early-mover advantages stem from the unique and valuable resources of being first, while the sustainability of early-mover advantages depends on the rarity and inimitability of these resources, while imitators struggle to catch up with early-movers (Barney, 1991). These advantages, in theory, can outweigh early-mover costs and create positive performance impacts (Lieberman and Montgomery, 1988). Consequently, the following conceptual framework not only centres around the questions on the imitation-performance relationship but also links it

to the inherent early-mover discussion, where early-movers can experience benefits while bearing certain costs imitators might not face.

2.3 Conceptual Framework

We draw on imitation and adjacent theories, such as absorptive capacity, diffusion, knowledge spillovers, mimetic isomorphism, vicarious learning, economic herding and early-mover discussion to develop our arguments and explore the imitation-performance relationship. Strategic imitation refers to the "attempt to reproduce [...] another firm's product, processes, capabilities, technologies or [strategic] decision [..]" (Posen et al., 2023, p 76). While the imitation-performance relationship is to a significant extent empirically unexplored, management research further developed into two very different theoretical schools of thought; on the one hand, implying a static understanding with negative performance effects and, on the other hand, a dynamic understanding with expected positive outcomes (Posen et al., 2023). The static model aligns with the RBV and the dynamic model with the dynamic capability literature, while imitation as a concept is furthermore linked and overlapping with a multitude of theoretical concepts within management research (Posen et al., 2023) and economics, suggesting an aggregate performance outcome lens. We furthermore extend our conceptual framework to account for early-mover effects (Hoppe, 2000; Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). Our conceptual framework is depicted in Figure 2.1. An operationalised and detailed description of the framework, including all variables, is further explained in the empirical study section (3).

To assess the performance-imitation relationship and be able to capture both fundamentally different outcome expectations of imitation, we employ a firm-centric macro-level lens and investigate if imitation on an aggregated level leads to negative or positive performance outcomes. We investigate whether firms imitate in relation to the market and, in turn, if this imitation leads to positive or negative performance outcomes. To account for the described theoretical frictions and consequently understand the aggregated financial performance outcomes of imitation, we borrow from the broader and widely accepted measure from (financial) herding scholars and transpose it into our strategic firm-level: We measure imitation as the observable correlated behaviour of a group of organisations, occurring consciously or unconsciously, as they pursue similar opportunities or evaluate topics and trends in similar manners at the same time.[3] The aggregated outcome measurement aligns with the few previous imitation-performance studies, which

[3]The measure of imitation as correlated behaviour as in the financial herding literature results form the lack of consistent empirical measurement in the management literature.The concrete mathematical measurement is explained in the methodological section of the thesis.

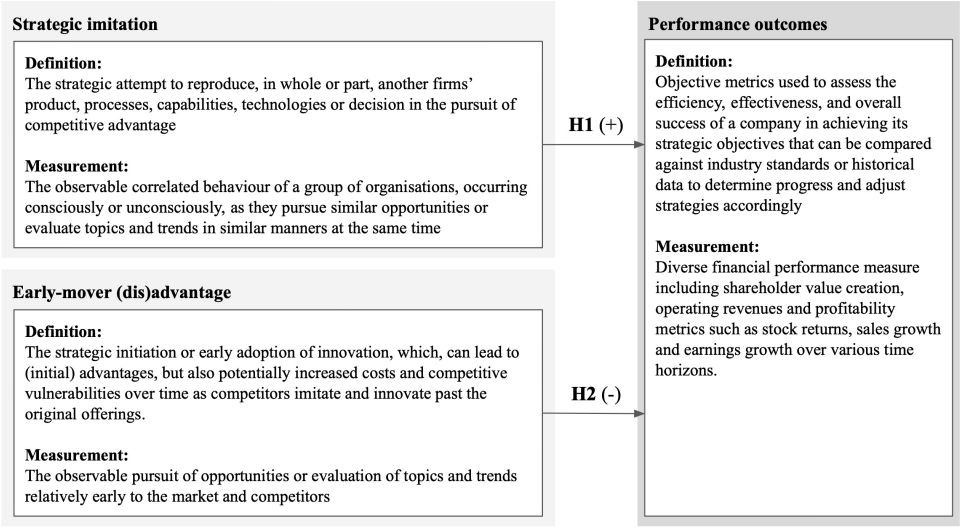

Figure 2.1. Conceptual Framework

also explored outcomes on an industry level (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999).

Building on (1) the dynamic capability perspective, (2) economic herding theories and (3) our aggregated imitation lens, we expect a positive imitation-performance relationship.[4]

First, from a dynamic capability perspective, imitation is not merely a reactive strategy but a deliberate enhancement of a firm's adaptive capabilities. Dynamic capabilities, as defined by Teece et al. (1997), include the ability to sense, seize, and transform resources effectively. In this context, strategic imitation can serve as a mechanism for firms to acquire and integrate external innovations that complement their existing capabilities, thereby fostering an environment of continuous learning and adaptation on an aggregated market level (Cohen and Levinthal, 1990; Levitt and March, 1988). By purposefully imitating, firms do not just copy but intelligently integrate these practices with their unique capabilities, leading to innovative combinations (Posen et al., 2023). This integration can enhance their responsiveness to market changes and technological advancements, leading to improved performance outcomes (Eisenhardt and Martin, 2000; Posen et al., 2023). Such an approach aligns with the dynamic capabilities framework, where the emphasis is not on resources alone but on the firm's ability to dynamically reconfigure existing assets and capabilities in response to external opportunities and threats.

[4]We assume a linear relationship between early-mover effects and performance, rather than a non-linear or e.g. inverted U-shaped one. This assumption is supported in the methodological section, where we test and confirm the standard multiple linear regression (MLR) assumptions.

Second, integrating economic herding theories offers further insights into why imitation might positively influence performance. Economic herding, often seen negatively due to its association with irrational conformity, can also be interpreted as a rational response under conditions of uncertainty and incomplete information (Banerjee, 1992; Bikhchandani et al., 1992). When firms observe a successful strategy or innovation adopted by peers, imitating such behaviours can be a strategic decision rather than a blind follow-the-crowd mentality. This strategic imitation might allow firms to minimise risks and costs associated with the development of untested strategies while capitalising on the proven success of others. Moreover, herding can lead to the standardisation of best practices across industries, creating efficiencies and improving the overall market performance. Consequently, in the dynamic model and according to herding theories, a firm would not converge to the individual performance maximum, but positive market effects, despite the suboptimal equilibrium, might still be more beneficial than relying solely on private information or the information of a few dyadic imitation relationships. By approaching imitation strategically, firms aim to navigate the search space and actively decide what information to consider. Thus, strategic herders or imitators combine their private with public information and potentially achieve better outcomes than relying solely on private or public information. Thus, when firms imitate based on strategic choices rather than blind herding, they are likely to see performance gains due to both reduced risks and increased adoption of successful innovations. Lastly, we do not expect a negative performance relationship as partially evident in financial herding when multiple investors pursue the same opportunity, and excess returns are shared among all investors (Jiang and Verardo, 2013). Our assumption that fractional returns do not decrease with the amount of additional investors is grounded in the strategic context of our study, where growth is not limited by financial market liquidity.

Third, the use of an aggregated lens combined with an intentional imitation process further supports the expected positive imitation-performance relationship. By viewing imitation not only as a dyadic interaction between the imitator and the imitated but as a broader market-level process, firms can discern patterns, absorb knowledge, and trends that signal successful strategies (Cohen and Levinthal, 1990; DiMaggio, Powell, et al., 1983; Jaffe, Trajtenberg, and Fogarty, 2000; Levitt and March, 1988; Strang and Soule, 1998). This macro-perspective allows firms to make more informed choices about what and when to imitate, turning imitation into a strategic tool that leverages collective industry insights. Furthermore, when firms intentionally choose what aspects of a competitor's strategy to imitate, they can avoid the pitfalls of wholesale imitation and instead focus on elements that complement their strategic goals and capabilities. This selective and intentional approach to imitation can lead to a more nuanced and effective adaptation of successful external strategies, ultimately enhancing firm performance by aligning imitation activities with specific organisational needs and market conditions.

In conclusion, a positive relationship between imitation and performance is supported by the integration of dynamic capabilities that enhance adaptive strategies, economic reasoning that justifies strategic conformity under uncertainty and improved search outcomes, and a macro-level intentional approach to imitation that leverages industry-wide insights for a strategic advantage.

Hypothesis 1 (H1): *A higher degree of imitation behaviour has a positive effect on firm performance*

Lastly, we extend our conceptual framework to account for early-mover effects. Early-mover effects refer to the strategic benefits and drawbacks of being among the first to enter a market or adopt a new technology (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). From the RBV, the advantages of being an early mover arise from possessing unique and valuable resources first. The sustainability of these early-mover advantages relies on the rarity and difficulty of imitating these resources (Barney, 1991). In theory, these advantages can surpass the costs of being an early mover, leading to positive performance outcomes (Lieberman and Montgomery, 1988). On the other hand, the dynamic capabilities framework suggests that the real strength of competitive advantages lies in the ability to adapt, integrate, and reconfigure internal and external competencies to address rapidly changing environments (Teece et al., 1997). Thus, dynamic capabilities enable a second-mover firm not just to keep up with early-movers but to continue evolving in ways that create new competitive advantages over time (Teece et al., 1997). This perspective is crucial in highly volatile industries where technological changes and market preferences shift swiftly (Eisenhardt and Martin, 2000).

Thus, building on (1) the costs and uncertainties faced by early-movers, which can erode their initial advantages proposed by the RBV, (2) the dynamic capabilities perspective that allows later entrants to efficiently assimilate and adapt using insights from early-movers, and (3) the critical role of knowledge spillovers in diminishing early-mover benefits, we expect a negative relationship between early-movers and their respective performance outcomes.[5]

First, while early-movers may initially seize substantial market share and secure key resources, they also face the burden of uncertainty and the high costs of pioneering new markets or technologies (Lieberman and Montgomery, 1988). Early-movers often bear the brunt of educating the market, refining immature technologies, and overcoming initial

[5]We assume a linear relationship between early-mover effects and performance rather than a non-linear, e.g. inverted U-shaped one. This assumption is supported in the methodological section, where we test and confirm the standard MLR assumptions.

market resistance—costs that later entrants can avoid (Lieberman and Montgomery, 1988). This scenario suggests that while early-movers initially establish a competitive stance, the costs and risks associated with these pioneering moves can erode their advantages over time as imitators learn from the early-movers' experiences and enter the market with fewer uncertainties and lower costs.

Second, the dynamic capabilities perspective emphasises the ability of firms to rapidly assimilate and adapt to new information and technologies (Teece et al., 1997). In this context, later entrants can leverage the trials and errors of early-movers, incorporating this knowledge to sidestep pitfalls and enhance their own strategic positioning (Eisenhardt and Martin, 2000; Suarez and Lanzolla, 2005). Thus, second-movers or later entrants, equipped with insights gained from observing early-movers, can deploy their resources more efficiently and effectively. This strategic imitation, coupled with dynamic capabilities, enables them to outmanoeuvre early-movers who might have exhausted substantial resources in initial explorations and may be less agile in responding to evolving market demands.

Third, the influence of knowledge spillovers plays a critical role in diminishing the benefits that early-movers might accrue (Jaffe, Trajtenberg, and Henderson, 1993; Jaffe, Trajtenberg, and Fogarty, 2000). As early-movers innovate and break new ground, they inadvertently share valuable insights through various channels—industry reports, patents, executive speeches, and more—which later entrants can exploit without incurring the initial costs of innovation (Hoppe, 2000; Jaffe, Trajtenberg, and Henderson, 1993). This involuntary transfer of knowledge can accelerate the capabilities of later entrants, allowing them to catch up or even surpass early-movers in terms of technological proficiency and market share. In consequence, early-movers again bear higher costs but experience in aggregation similar benefits as their learning peers.

Therefore, the hypothesis that being an early-mover negatively affects firm performance is supported by the realisation that early-mover advantages are not only temporary but are also vulnerable to rapid erosion through the strategic actions of dynamic, strategically imitating competitors. These competitors effectively utilise the lessons learned from early-movers, thereby diminishing the supposed 'first-mover advantage' and potentially turning it into a disadvantage as market conditions evolve. In summary, the cumulative costs and strategic responses from imitators contribute to market dynamics, where early movers might see their initial benefits reversed over time caused by their imitating peers.

Hypothesis 2 (H2): *Being an early-mover has a negative effect on firm performance*

3

Empirical Study

This section contains (1) a general and concrete introduction to our empirical study, followed by a detailed description of (2) the data and context, (3) the employed variables, (4) the respective descriptive statistics, correlations, and model-free evidence, (5) the panel regression models, (6) the study's main results and (7) the validation approach.

3.1 Research Design

Generally, our study employs a positivist (Straub and Gefen, 2004) and deductive approach (Gill and Johnson, 2010): We analyse the imitation-performance relationship by developing and testing hypotheses based on established theories and empirical evidence (Bell et al., 2022; Edmondson and Mcmanus, 2007). This approach is common within the imitation-performance literature (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022). We deductively build our hypotheses on the dynamic capability literature and economic herding theories (Banerjee, 1992; Barney, 1991; Bikhchandani et al., 1992; Peteraf, 1993; Teece et al., 1997) to resolve the nascent friction within the static and dynamic understanding of imitation (Posen et al., 2023). We introduce large-scale machine-learning enabled LDA and sentiment-based independent variables (Bayer et al., 2017; Bellstam et al., 2021; Blei et al., 2003; Li et al., 2013; Ligorio et al., 2022; Miric et al., 2023). We use objective and diverse performance measures such as stock returns, revenue growth and profitability growth (Berk and DeMarzo, 2007; Eberhart et al., 2004). For the analysis, we use statistically robust pooled, fixed-effects and random-effects panel regressions (Berry, 1993; Torres-Reyna, 2007; Wolf and Best, 2013). This decision is based on the objective to overcome frictions of expected performance outcomes between established management theories (Miller and McKee, 2021; Samba et al., 2021).

Concretely, we study the imitation-performance relationship within the context of innovation, analysing corporate sentiment with more than 5000 observations from S&P 500 firms during 2009-2023.[1]

Innovation is theoretically woven into the imitation literature and is characterised by the utilisation of advanced technologies to introduce new products, services, business processes, or business models that improve performance (Christensen, 1997; Nguyen and Agrawal, 2023; Zhou, Yim, et al., 2005). Theoretically, imitation has been discussed together with innovation within the RBV (Barney, 1991; Peteraf, 1993) and since then studied throughout the strategy literature in an ongoing symbiotic relationship (Posen et al., 2023). Nguyen and Agrawal (2023) contribute to the innovation-performance field with their DIS dictionary, which includes the 500 most disruptive innovation-related keywords, validated by their robust financial analysis and proven relevance. This disruptive innovation dictionary provides the initial foundational context for our study. Yet, to address potential limitations of pre-existing dictionaries, imperfect contextualisation of innovation topics, and changes over time (Miric et al., 2023), we developed our own innovation dictionary across time using LDA (Blei et al., 2003). LDA topic modelling is a type of statistical machine learning model used in natural language processing and text mining (Blei et al., 2003). It is a generative probabilistic model that explains similar sets of words within a document by unobserved groups (Blei et al., 2003). This approach aligns closely with our specific conceptual framework and captures changes in the innovation context itself, which is crucial, as ongoing change can limit the time window available to undertake and implement imitation strategies (Posen and Levinthal, 2012), and furthermore opens the door to explore additional contexts, thereby broadening the scope of our analysis of the imitation-performance relationship and the potential general theoretical contributions. Concretely, we employ textual analysis to explore the under-researched imitation-performance relationship (Posen et al., 2023), as it is already a common practice in the closely related innovation-performance literature (Bellstam et al., 2021; Nguyen and Agrawal, 2023; Nousiainen et al., 2022).

Textual analysis allows us to measure imitation as the observable correlated behaviour of a group of organisations, occurring consciously or unconsciously, as they pursue similar opportunities or evaluate topics and trends in similar manners at the same time. We measure imitation as the correlation of communication indicators in the annual reports of the observed companies with the wider market. The observed communication indicators are e.g. the mentions and sentiment towards the dictionary-defined context.[2] Textual analysis

[1] Each of the 5000 observations consists of six data points for each variable and topic. See Table 3.1 for an overview of all the imitation and early-mover variables.

[2] See Table 3.1 for a description of the communication indicators including mentions, sentiment and sentiment breadth.

of annual reports (10-K reports) of companies is a common practice within a multitude of management research from innovation to competition and marketing studies (Bayer et al., 2017; Bellstam et al., 2021; Li et al., 2013; Nousiainen et al., 2022). The 10-K report, mandated annually by the U.S. Securities and Exchange Commission (SEC), offers a comprehensive overview of the company's business, including technology, operations, financials, strategy, and future outlook. We follow the commonly used reasoning that a company involved in groundbreaking innovations is compelled to disclose this information in the report due to its role in informing current and potential investors about the firm's true state and prospects (e.g., Nguyen and Agrawal, 2023). SEC filings are an important means of corporate communication that are required to comply with regulatory standards emphasising transparency and accuracy. These standards aim to ensure that the information provided is not solely promotional but offers a truthful and comprehensive representation of the company's activities and state, helping to protect investors from potentially overly optimistic or misleading claims(Nguyen and Agrawal, 2023).[3]

Following Posen et al.'s (2023) call for investigation of the imitation-performance relationship with robust and diverse performance measures, we test for shareholder returns and operating performance such as revenue growth and profitability growth. These objective outcome measures help overcome a common pitfall in management research where theoretical frictions are difficult to resolve because studies predominantly rely on subjective outcome measures such as surveying or interviewing the management team or conducting small-sample case studies (Miller and McKee, 2021; Samba et al., 2021). Our DIS innovation context starting point was historically able to predict stock returns and showed relevance for operational performance measures (Nguyen and Agrawal, 2023).

We summarise our empirical study derived from the conceptual framework in Figure 3.1. An overview of the collected data, as well as the extracted variables, are described in Figure 3.2 and listed in Table 3.1. The figure and table are explained in the subsequent text sections 3.2 and 3.3.

[3]We are aware that despite high regulatory standards, the 10-K might still be vulnerable to e.g. human errors and biases.

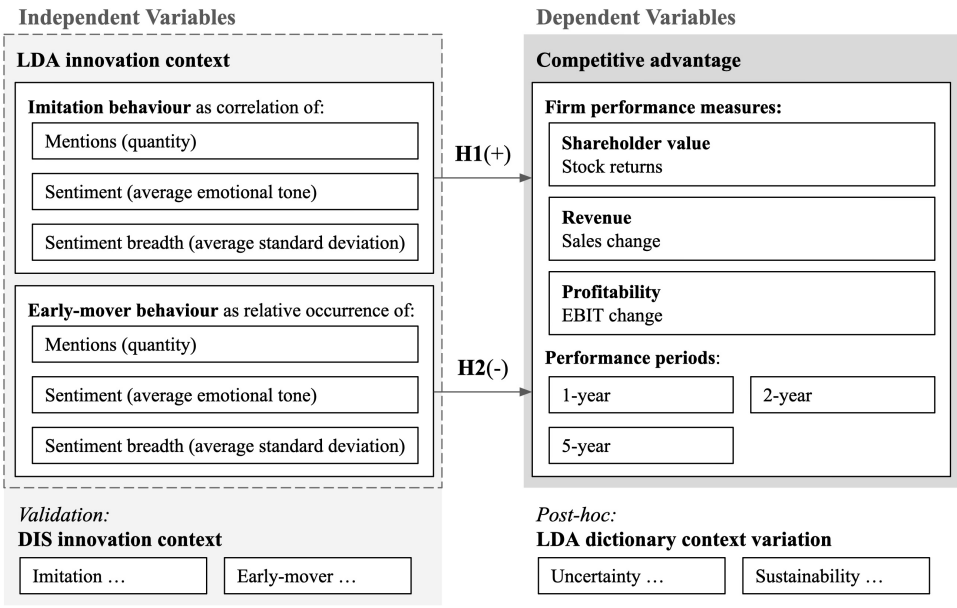

Figure 3.1. From Conceptual Framework to Empirical Model

3.2 Data & Context

Our study's dataset comprises the annual reports and respective performance outcomes of companies listed in the S&P 500 index during the observation period from 2009 to 2023. S&P 500 companies operate globally and represent diverse industries while demonstrating dynamic capabilities essential for maintaining their index position (*SP 500® Equity* 2024; Uotila et al., 2009). To collect the sample, we systematically scrape the S&P 500 index to capture the evolving company composition. We aim to mitigate survivorship bias (Fung and Hsieh, 1997) by continually tracking drop-outs and new add-ins of companies into the index. Annual 10-K reports, mandated by the SEC and sourced from the University of Notre Dame SEC/Electronic Data Gathering, Analysis, and Retrieval (EDGAR) database, serve as the basis to later define our innovation context and offer comprehensive insights into strategic imitation behaviour. 10-K reports are an established data source in innovation studies (Bayer et al., 2017; Bellstam et al., 2021; Li et al., 2013; Nousiainen et al., 2022). Performance metrics sourced from the WRDS Center for Research in Security Prices, LLC (CRSP)/Compustat Merged databases cover robust and diverse financial and operating metrics from 2009 to 2023, providing insights into shareholder value creation (stock market returns), operating revenues (sales) and profitability metrics (Earnings before interest and taxes, EBIT). Diverse and objective performance measures mitigate issues with subjective performance measures (Miller and McKee, 2021; Samba

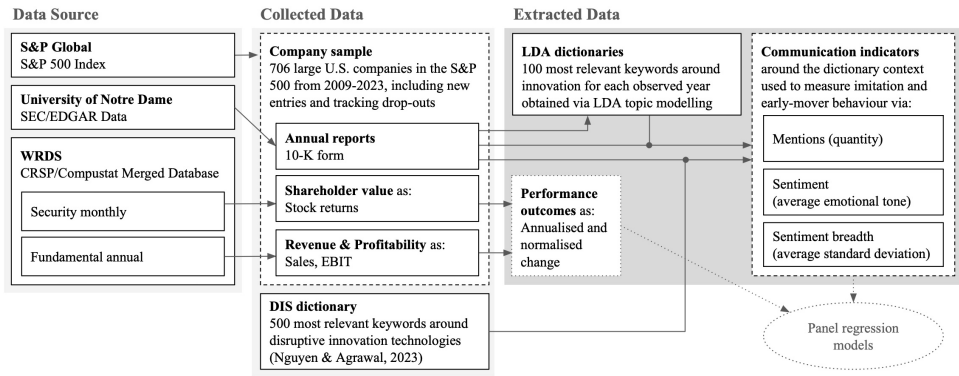

Figure 3.2. Data Collection and Extraction

et al., 2021) and enhance the study's robustness by overcoming previous limitations within the imitation-performance literature (Posen et al., 2023). The data flow, including the respective sources, collection and extraction, is summarised in Figure 3.2 and explained in detail in the following subsections.

3.2.1 Company Sample

Our initial company dataset consists of 706 companies listed in the S&P 500 index during the years 2009 to 2022. The S&P 500 is a common sample size for text-based research in an innovation context (Bellstam et al., 2021). The S&P 500 index is created by the company S&P Global and offers a broad perspective on corporate America by tracking approximately 80% of the available market capitalisation of listed U.S. corporations (*SP 500® Equity* 2024). Despite comprising only American companies, the organisations within the index operate globally, with revenues created from across the world (*SP 500® Equity* 2024). The companies in the index furthermore reflect a diverse range of industries and economic sectors (*SP 500® Equity* 2024). Moreover, the dynamic capabilities of these companies are implicit since they must continuously adapt and leverage both explorative and exploitative strategies to maintain their position in this market capitalisation-based index (Uotila et al., 2009). The ability of companies to adapt highlights the relevance of our company sample to examine the imitation-performance-relationship amidst viewing imitation as a potential enabler of dynamic capabilities rather than merely an erodent of competitive advantage, as suggested by the RBV.

We argue that the sample selection period 2009-2022 is suitable for investigating the imitation-performance relationship by covering the recent decade of economic growth starting after the highly volatile financial crisis. Furthermore, the combination of the large sample size and selection of recent years overcomes another limitation of the few existing

empirical studies on the imitation-performance relationship: previous studies contained limited and single-sector data mainly from the 1980s, 1990s, and early 2000s (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022).

We collect our sample of the S&P 500 companies by systematically scraping the index composition at six-month intervals from 2009 to 2022. Our sample includes a company if it was part of the SP 500 at any time during this period. This approach aims to capture the continuously evolving index composition and ensure that our company sample includes not only companies that remained within the S&P 500 throughout the study period but also those firms that were removed or added. By doing so, we aim to mitigate survivorship bias, which could skew insights if only companies that remained in the index by 2022 were considered. Survivorship bias can be problematic in empirical studies with financial performance outcomes, as excluding companies that drop out of the index could lead to an overly optimistic portrayal of corporate health or market dynamics (Fung and Hsieh, 1997). We do not include companies that first appeared in the index in 2023, as we aim to capture at least one year of financial performance after the respective observation period. After scraping the index composition, we aggregate the collected company sample into a single database, grouped by each company's ticker symbol. This step is essential to address instances where companies miss Central Index Key (CIK) values and ensure accurate identification in subsequent regulatory 10-K filings. We furthermore remove companies with insufficient observation periods. An observation period refers to the period in which we observe a company's imitation and early-mover behaviour. The minimum observation period is five years; thus, a company with sufficient data published at least five years of consecutive annual reports during 2009-2022. Furthermore, a company must have at least one year of financial performance data for at least one performance variable available after the sufficient observation period. Depending on data availability, respective performance variables, and outlier handling, the final company sample size per regression ranges from 590 to 612 observed entities. Later sections of this study describe the specific observation measures, performance variables, and the outlier handling procedure.

3.2.2 Annual Reports

We analyse 9552 annual reports from the S&P 500 companies listed in the index during 2009 and 2022. The annual reports are sourced from the University of Notre Dame SEC/EDGAR database. The annual report, or 10-K form, is a comprehensive document annually mandated by the SEC and offers an in-depth disclosure of a company's financial results, business model, and the potential risks and opportunities it foresees. This form provides a level of detail and scrutiny surpassing that typically found in shorter annual reports that companies distribute on their websites or during shareholder meetings. Consequently, the 10-K report is not only pivotal for communicating with current and potential

investors but must also adhere to strict SEC regulations that demand comprehensive transparency and accuracy. This regulatory environment supports the assumption that the annual reports serve as a reliable tool for understanding corporate states and prospects while aiming to protect investors from potentially misleading overstatements (Nguyen and Agrawal, 2023).

Analysing firm sentiment through the lens of the annual 10-K report is a well-established practice in management research, spanning areas from innovation to competitive strategy and marketing (Bayer et al., 2017; Bellstam et al., 2021; Li et al., 2013; Nousiainen et al., 2022). Consequently, the decision to empirically explore imitation behaviour in an innovation context is based on the wide acceptance in management research and regulatory requirements. The regulatory requirements mandate that companies shall be reliable and precise when, for example, expressing their sentiment about a given topic within our context dictionaries.

The annual reports serve as the basis to create both the innovation context dictionary via LDA topic modelling and to directly analyse the imitation behaviour. LDA topic modelling is a type of statistical machine learning model used in natural language processing and text mining (Blei et al., 2003), and both the context creation and the imitation behaviour variables are explained in detail in Section 3.3.

3.2.3 Performance Data

Our performance data consists of the diverse financial performance of each S&P 500 company from 2009-2023. The performance data includes security performance in the form of stock returns and operation performance in the form of revenue changes and EBIT changes. The raw monthly security performance was queried by the CIK from the Security Monthly WRDS CRSP/Compustat Merged database. Similarly, the raw operating performance data were sourced from the Fundamentals Annual WRDS CRSP/Compustat Merged database. Concretely, the stock returns reflect the percentage change in stock price over a given period (Berk and DeMarzo, 2007). Sales measure the total revenue a company generates (Berk and DeMarzo, 2007). EBIT, or Earnings Before Interest and Taxes, highlights a company's profitability, excluding the impact of structural financing decisions and their tax impacts (Berk and DeMarzo, 2007). Sales and EBIT change are then percentage decreases or increases of sales and EBIT over a certain period. In general, the financial market and operating performance variables are common outcome measures in management research and innovation contexts (Bellstam et al., 2021; Eberhart et al., 2004) as well as in combination with text-based and 10-K-based independent variables (Bayer et al., 2017; Li et al., 2013; Nousiainen et al., 2022).

A limitation of the stock return measure is that hundreds of factors can influence stock performance (Harvey et al., 2016). We are partly mitigating this limitation by enhancing the robustness through the described operating performance metrics in the form of sales change and profit change, which is closer to the fundamental core of the observed organisations (Berk and DeMarzo, 2007). However, a common critique of accounting-based operating metrics is that accounting practices allow companies to report and design their operating results more or less conservatively and use them as a strategic communication tool itself (Robinson et al., 2012). Nevertheless, the combination of shareholder value (stock returns), revenue (sales change) and profitability (EBIT change) data over various time horizons not only aligns with Posen et al.'s (2023) call for robust and diverse outcome measures to explore the imitation-performance relationship but helps to overcome a general pitfall in management research where theoretical frictions are difficult to be resolved because studies predominantly relied on subjective outcome measures such as surveys (Miller and McKee, 2021; Samba et al., 2021).

3.3 Variables

This subsection contains a description of all variables used in our empirical model and how they are extracted from the empirical data. We describe (1) the context dictionaries based on the annual reports and (2) the communication indicators extracted from the annual reports based on the context dictionaries. We furthermore explain the independent variables such as (3) the imitation variables, (4) the early-mover variables as well as (5) the performance variables serving as dependent variables. All variables are summarised in Table 3.1 and explained in the following subsections.

Table 3.1. Measures, Sources, and Supporting Literature of Used Variables

Variable	Measure	Source	Supporting literature
Dictionary variables LDA created context variables that define and delimit the context of the study.			
LDA innovation	Main context defined by LDA-created dictionaries containing the 100 most important words related to innovation from 2009-2022	This study, using LDA topic modelling on 10-K forms from the University of Notre Dame SEC/EDGAR database	The current study Bayer et al., 2017 Nousiainen et al., 2022

Continued on next page

Table 3.1 – continued from previous page

Variable	Measure	Source	Supporting literature
DIS innovation	Dictionary of the 500 most important words in the disruptive innovation context for robustness checks and validation of the main context	Adjacent paper using a semi-supervised machine learning approach, focusing on 29 disruptive technologies	Nguyen and Agrawal, 2023 Bloom et al., 2021
LDA sustainability	Post-hoc context defined by LDA-created dictionaries containing the 100 most important words related to sustainability from 2009-2022	This study, using LDA topic modelling on 10-K forms from the University of Notre Dame SEC/EDGAR database	The current study Bayer et al., 2017 Ligorio et al., 2022
LDA uncertainty	Post-hoc context defined by LDA-created dictionaries containing the 100 most important words related to uncertainty from 2009-2022	This study, using LDA topic modelling on 10-K forms from the University of Notre Dame SEC/EDGAR database	The current study Bayer et al., 2017
Imitation variables Independent variables to capture the imitation behaviour of the individual firms for all 5-year observation windows during the period of 2009-2022. Variables are based on the underlying communication indicators.			
Imitation mentions (also Im. mentions)	Correlation of a company's topic mentions with the aggregated market within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study Giachetti, Lampel, et al., 2017 Giachetti and Pira, 2022 Li et al., 2013

Continued on next page

Table 3.1 – continued from previous page

Variable	Measure	Source	Supporting literature
Imitation sentiment (also Im. sentiment)	Correlation of a company's sentiment on a topic with the aggregated market sentiment within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study
Imitation sentiment breadth (also Im. sen. breadth)	Correlation of sentiment standard deviation on a topic by a company with the aggregated market within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study
Early-mover variables Independent variables to capture early-mover behaviour for all 5-year observation windows during the period of 2009-2022. Variables are based on the underlying communication indicators.			
Early-mover mentions (also E-m. mentions)	Above-average topic mentions by a company compared to the aggregated market within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study Lieberman and Montgomery, 1988

Continued on next page

Table 3.1 – continued from previous page

Variable	Measure	Source	Supporting literature
Early-mover sentiment (also E-m. sentiment)	Above-average sentiment by a company on a topic compared to the aggregated market within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study Lieberman and Montgomery, 1988
Early-mover sentiment breadth (also E-m. sen. breadth)	Above-average sentiment standard deviation by a company compared to the aggregated market within a given context	This study, via computation based on data from LDA context dictionaries and annual reports (10-K forms) from the University of Notre Dame SEC/EDGAR database	The current study Lieberman and Montgomery, 1988
Performance variables Dependent variables to capture firm performance outcomes for 1-year, 2-year, and 5-year periods during 2009-2023.			
Stock returns	Financial market performance, reflecting shareholder value creation	WRDS CRSP/Compustat Merged Database - Monthly Security Data	Posen et al., 2023 Bellstam et al., 2021 Eberhart et al., 2004 Li et al., 2013 Bayer et al., 2017 Nousiainen et al., 2022
Sales change	Change in revenue, reflecting operating performance	WRDS CRSP/Compustat Merged Database - Annual Fundamental Data	Posen et al., 2023 Bellstam et al., 2021 Giachetti and Pira, 2022
EBIT change	Change in earnings before interest and taxes (EBIT), reflecting profitability	WRDS CRSP/Compustat Merged Database - Annual Fundamental Data	Posen et al., 2023 Eberhart et al., 2004

3.3.1 Dictionaries

This subsection centres around the creation of the dictionaries used to define or study context. We describe (1) the purpose of the dictionaries, (2) the description of the LDA topic modelling, including a reflection on the potential strengths and weaknesses of this method, and (3) the contextual starting point and reasoning to employ LDA topic modelling.

First, to quantitatively analyse the aggregated imitation behaviour in the large dataset of annual reports, we need dictionaries containing the keywords describing the desired context of observation. The keywords within these dictionaries then serve as the basis to analyse the mentions and sentiment of the dictionary keywords to describe how each firm communicates around the concept or set of ideas within the dictionary context. Consequently, the definition of the dictionaries, and thus the context of the study, is an essential step of our research design (Bayer et al., 2017; Ligorio et al., 2022; Miric et al., 2023; Nousiainen et al., 2022).

Second, LDA topic modelling is a machine learning technique that can be used to automatically extract topics within a set of documents (Blei et al., 2003), in this case, the annual reports. This method has been used in previous management studies and is an emergent method to analyse text data (Bayer et al., 2017; Ligorio et al., 2022; Nousiainen et al., 2022). The results of LDA topic modelling are a set of generated topics that are described by the most likely words to co-occur within the same context. While the topics themselves are created by an unsupervised machine learning model, they can afterwards be interpreted and selected by a human depending on the research context (Tirunillai and Tellis, 2014).

Third, Nguyen and Agrawal's (2023) DIS dictionary with the 500 most disruptive innovation-related keywords serves as a starting point and later validation context of our study. However while doing so, we furthermore address potential limitations of pre-existing dictionaries, imperfect contextualisation of innovation topics and changes over time (Miric et al., 2023), by developing our own innovation dictionary across time using the machine-learning technique LDA topic modelling (Blei et al., 2003). The DIS dictionary, initially built on 29 disruptive technologies identified by Bloom et al. (2021) in a semi-supervised machine learning-based textual analysis of Initial Public Offering (IPO) prospectuses, obtains the 500 most relevant keywords around those technologies with the goal of being less dependent on patent data (Nguyen and Agrawal, 2023). However, as the aim of this study is to analyse how companies react and communicate around wider trends and measure the aggregated imitation behaviour, appropriate dictionaries are required that adequately describe these innovations and account for changes in the language and relevance of certain keywords (Miric et al., 2023). Consequently, employing LDA

topic modelling on the annual reports aligns more closely with our specific conceptual framework variables, captures changes in the innovation context itself, but also opens the door to exploring additional contexts, thereby broadening the scope of our analysis of the imitation-performance relationship and the potential theoretical implications. The DIS dictionary was used as a benchmark and identification reference for the dictionaries developed as part of this study. Keeping the DIS dictionary furthermore offers the opportunity to use it as potential validation of the new LDA innovation dictionaries developed by this study.

Based on our DIS starting dictionary, we identified the innovation dictionary as the primary lens, defining the study context. The LDA approach allowed us to identify further topic dictionaries, such as uncertainty and sustainability, serving as the basis for the post-hoc analysis. Since topics remain relatively stable over time, we could track these dictionaries throughout our observation period. Tracking yearly dictionaries allows us to capture how companies' communication around selected topics evolves. The ability to use words describing innovation accurately in both 2009 and 2022 is a significant and novel advantage of this method. It ensures we capture the changing innovation terminology over the 13-year observation period. By using the relevant dictionary for each year, we also avoid the ex-ante survivorship bias of, e.g. applying the latest dictionary from 2022 to analyse the annual reports from 2009.

On the other hand, while LDA topic modelling can address the limitations of manual dictionary creation and relevance issues related to timing (Miric et al., 2023), by automatically identifying words commonly used, the output often results in a lengthy list of topics with associated words (Blei et al., 2003). These words may not be universally useful or interpretable. Consequently, it is necessary to select topics based on their relevance and coherence within the broader context. This selection process introduces potential bias, as it relies on human judgement (Tirunillai and Tellis, 2014). The chosen dictionaries must include words that capture broader trends relevant to all companies within a market. Based on our DIS selection lens, we can confidently observe that the words in the LDA innovation dictionary effectively capture how companies discuss relevant topics.

All used LDA dictionaries are listed in the Appendix . We can already preempt that the quality and relevance of findings from the later described empirical model support the chosen LDA approach with updated dictionaries over the entire observation period.

3.3.2 Communication Indicators

To quantitatively measure how companies strategically communicate around the defined context and topics of the dictionaries and consequently investigate imitation as the observable correlated behaviour of a group of organisations as they pursue similar opportunities

or evaluate topics and trends in similar manners at the same time we measure a firm's (1) mentions of a topic, (2) the sentiment or emotional tone around the topic and (3) the deviation within the respective sentiment.

First, the mentions or quantity counts how often words in the dictionary occur in the respective annual report. The mentions are not normalised according to the report length, as this could introduce statistical entity-specific effects depending on the respective reporting length of a company. The raw mentions inherently aim to describe the robust timing of when and how intensively a company mentions a topic. Capturing timing has furthermore been the key independent variable in previous empirical works on the imitation-performance relationship observing mobile phone vendors in the UK from 1997-2008 (Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022). Mentions are furthermore a common variable when investigating competition based on 10-K form textual data (Li et al., 2013).

Second, sentiment is defined as the emotional tone towards or average evaluation of a topic and whether the firm, e.g. perceives a technology as useful or not if they plan to use or oppose it (Davis, 1989; Devaraj et al., 2008; Schepman and Rodway, 2023). The sentiment variable is obtained by performing sentiment analysis on sentences containing one or several of the words within a context dictionary. The sentiment score is the output from a machine learning model that scores a sentence on its expression and whether it is positive or negative. The concrete model used is VADER (Valence Aware Dictionary for sEntiment Reasoning) developed by Hutto and Gilbert (2014) and considered as the "gold-standard of lexicon-based sentiment analysis for numerous types of textual content" (Al-Natour and Turetken, 2020, p 4; Ribeiro et al., 2016). The obtained sentiment score values range from -1 to 1, with negative values, zero and positive values, respectively, indicating a negative, neutral and positive sentiment. The average sentiment around a topic is then calculated over a company's entire annual report for the observed year. While exploring the imitation-performance relationship based on corporate sentiment is novel, the sentiment variable itself strongly aligns with our definition of imitation to capture the evaluation of a topic and serve as the proxy of a firm's attitude towards the explored contexts.

Third, sentiment breadth or the average sentiment standard deviation is a measure of how nuanced a firm's communication is. By using the same sentiment analysis, the sentiment breadth variable can be extracted by calculating the standard deviation of the previous raw sentiment variable throughout the entire text of each annual report. The standard deviation of the sentiment variable describes the amount of variation in the distribution. A low standard deviation characterises a company that is very consistent in discussing the topic of the dictionary, while a large standard deviation would describe a company that expresses a broader variety of sentiments around the dictionary contexts.

This extension of the sentiment communication indicator might be worthwhile to obtain a more accurate understanding of a company's evaluation of a topic. On the other hand, deviation within sentiment is a novel and abstract construct, and thus potentially less clear and less relevant than the pure sentiment and raw mentions count. Furthermore, the existing imitation-performance literature does not provide support for the sentiment breadth variable, suggesting to primarily evaluate the hypotheses on the more robust mentions and sentiment variables. To align the communication indicator and measurement approach with our mature research state (Edmondson and Mcmanus, 2007) the sentiment breadth measure should only be used in a reflected manner and potentially serve as a extended starting point for potential theoretical discussion.

The analysis and extraction of the communication indicators are performed on all annual reports for all concerned companies and within the time range of the study. This information can then be used to assess how a company's communication changes over the years and how the aggregated S&P 500 communication around the respective contexts evolves. This results in a communication data set of approximately 5000 data points for each context with six communication indicators included in a data point. The exact number of utilised observations furthermore depends on the year of observation and performance window.

Lastly, the individual communication indicators are utilised not only to measure imitation behaviour as a correlation with the market but also to test the early-mover hypothesis. The early-mover variables are derived from the respective market average for each communication indicator. The following subsections explain both the imitation and early-mover variables, outlining how they are calculated based on the individual communication indicators.

3.3.3 Imitation Variable

Based on our research questions to investigate the imitation-performance relation and measure the imitation as the observable correlated behaviour of a group of organisations, occurring consciously or unconsciously, as they pursue similar opportunities or evaluate topics and trends in similar manners at the same time, we need a quantitative variable that captures if a company's behaviour moves along with the aggregated market average, or independently of its peers. Here, the communication indicators serve as a proxy for their accurate behaviour, as mandated by the SEC. Therefore, the chosen imitation measurement can be defined as the correlation between a company's communication indicators and the market average of the same communication indicator along the time axis.

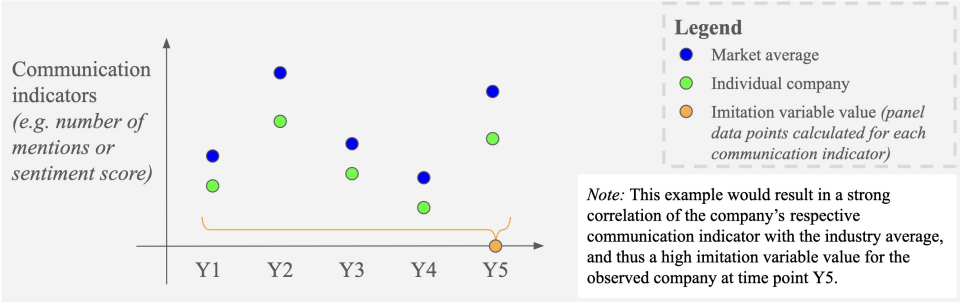

Figure 3.3. Imitation Variable

Here, a correlation describes a statistical relationship between two variables. The concrete type of correlation metric is Spearman's rank correlation coefficient. A value of 1 and -1 describes a perfect monotone relationship, respectively positive or negative. The correlation is then used as a quantitative variable for imitation, assuming that if a company's communication indicators around a specific topic are highly correlated to the S&P 500's communication indicators, imitation behaviour could be at play. Each of the company's communication indicators is compared to the average over the entire S&P 500 as depicted in Figure 3.3.

For example, a valid data point for performing correlation analysis involves a company that has consistently published annual reports throughout a relevant observation period. We extract communication indicators from these reports and calculate the correlation coefficient between these company-specific values and the aggregated market communication indicators for the same observation period. A high correlation coefficient indicates that the company's communication patterns align closely with broader market trends.

Our imitation variable is designed to capture the correlation between industry-wide averages of various communication indicators and those specific to individual companies. This correlation tends to be more reliable over larger time frames, as smaller time windows may introduce more uncertainty due to a reduced number of data points, with a potentially higher degree of statistical noise (e.g. a 3-year observation window of the imitation behaviour would be too short). However, extending the imitation window indiscriminately is also not feasible, as it would reduce the number of available data points for the subsequent performance analysis. A too-extensive approach restricts the number of usable data points and increases uncertainty and bias.

Therefore, we address this tension and select a sufficiently long 5-year imitation window to ensure robustness and maintain a higher number of data points, allowing

for a diverse range of starting and ending years for a more comprehensive imitation-performance analysis. We expect a robust balance for 5-year imitation windows allowing for multiple mid-term performance periods. For example, a moving 7-year window - consisting of 5-year imitation observation periods and 2-year performance observation periods - can be fit multiple times into our entire sample period from 2009-2023. An example of the combined observation periods is visualised in Figure 3.5.

3.3.4 Early-Mover Variable

We furthermore design a variable to capture early-movers. While the questions of how to define an early-mover allows for ambiguity and flexibility (Lieberman and Montgomery, 1988), our early mover variable centres around timing aspects (Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022) and aims to investigate the classical first-mover and second mover tensions (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). Our early-mover variable is calculated based on a company's above-average mentions, sentiment or breadth within an observation time window. It involves summing the differences between the company's communication indicators and the industry average for each year within the observation window as depicted in the formula below and visualised in Figure 3.4.

$$EarlyMover = \sum_{n=1}^{5} CompanyIndicator_n - IndustryAvg_n$$

A high early-mover variable indicates that a company communicated about a specific topic within the time window of interest more than others and does so either more frequently, more positively, or more nuanced, depending on the specific communication indicator used. This variable helps to identify companies that may have gained competitive advantages by addressing emerging trends and topics ahead of their peers. On the other

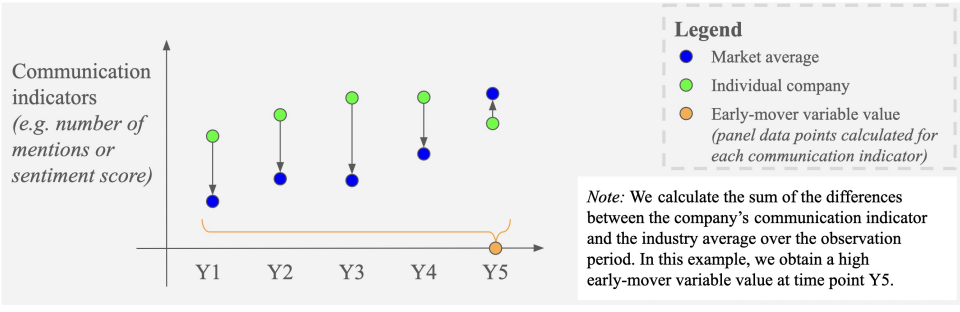

Figure 3.4. Early-mover Variable

hand, a potential weakness of our early-mover variable is the reliance on quantifying early-mover advantages directly through the communication indicators during the observation period and not during the observation-unrelated actual emergence of a topic which could have been already before 2009. Furthermore, and especially for the early-mover discussion, communication indicators might not accurately reflect the actual timing impact of actions taken by the company.

3.3.5 Performance Variables

The used performance variables are stock returns, sales change and EBIT change over various time periods (1-year, 2-year and 5-year). Stock returns reflect the created shareholder value, sales change covers the revenues, and EBIT the profitability of a firm (Berk and DeMarzo, 2007). These measures of robust and diverse financial performance outcomes across various time periods align with common practices in management research and our adjacent methodological decisions (Bellstam et al., 2021; Eberhart et al., 2004; Edmondson and Mcmanus, 2007). To construct our dependent variables, we annualise the monthly stock returns into one-, two-, and five-year performance windows. We furthermore compute the changes in sales and EBIT over one-, two-, and five-year periods, contingent upon the availability of sufficient historical data. The normalisation of performance measures such as sales into sales change has previously been used in innovation contexts in combination with textual analysis (Bellstam et al., 2021).

Previous empirical studies of the imitation-performance relationship furthermore used Return on Assets (ROA) as a performance outcome variable (Barreto and Baden-Fuller, 2006; Deephouse, 1999). ROA describes a firm's net income divided by aggregated assets (Berk and DeMarzo, 2007). While Barreto and Baden-Fuller (2006) and Deephouse (1999) both study the imitation-performance relationship within a single industry with very similar companies (banking), we acknowledge that business models and asset structures differ significantly across our company sample (Malone et al., 2006). Consequently, we do not include ROA as an outcome measure for our large-sample size study of the observed diverse companies. We rely on the normalised changes of the shareholder value creation, sales and EBIT to ensure comparability across firms (Bellstam et al., 2021). Firm-specific characteristics are furthermore accounted for by our statistical test and robust model specification as described in Section 3.5.

3.4 Descriptive Statistics

In the following section, we present (1) descriptive statistics to summarise the data collected and (2) model-free evidence as the raw correlation between the independent and dependent variables to demonstrate that similar conclusions can be reached through

simpler analytical approaches, and thus reduce the risk that results solely rely on specific model assumptions.

First, Table 3.2 and Table 3.3 provide a summary of the descriptive statistics for each independent and dependent variable, including the mean, standard deviation and distribution of values dissected into the minimum and maximum value as well as 25%, 50% and 75% percentiles. Table 3.4 describes the correlation between all independent imitation and early-mover variables, where, in general, most of the correlations are weak, and furthermore, all correlations between imitation and early-mover variables are very weak, supporting the robustness of our approach.

Table 3.2. Descriptive Statistics Independent Variables

Variable	mean	std	min	25%	50%	75%	max
Im. mentions	-0.0226	0.5003	-1.0000	-0.4000	0.0000	0.3000	1.0000
Im. sentiment	0.0123	0.5156	-1.0000	-0.4000	0.0000	0.4000	1.0000
Im. sen. breadth	0.2089	0.5551	-1.0000	-0.2000	0.3000	0.7000	1.0000
E.-m. mentions	123.7268	5255.1065	-9445.1566	-3394.6412	-965.9353	2371.4331	41068.8973
E.-m. sentiment	0.0025	0.2271	-1.1158	-0.1306	0.0164	0.1468	0.8286
E.-m. sen. breadth	-0.0046	0.1158	-0.4291	-0.0763	-0.0103	0.0594	0.6216

Table 3.3. Descriptive Statistics Dependent Variables

Variable	mean	std	min	25%	50%	75%	max
Returns 1y	16.638690	44.150193	-99.079457	-5.057410	14.266314	33.973606	1332.704403
Sales 1y	0.077097	0.244565	-0.903951	-0.005263	0.054272	0.130195	6.475100
EBIT 1y	0.045649	6.739431	-323.384615	-0.073613	0.061480	0.175458	303.185307

Note: This table only describes the statistical characteristics of the 1-year performance variables. All 2-year and 5-year performance variables are the aggregation of the respective 1-year performance data.

Table 3.4. Summary of Correlations Between Independent Variables

Variable	1	2	3	4	5	6
1 Im. mentions	1.000000					
2 Im. sentiment	0.084247	1.000000				
3 Im. sen. breadth	-0.164965	0.052884	1.000000			
4 E.-m. mentions	0.024760	0.012558	0.016383	1.000000		
5 E.-m. sentiment	0.031298	-0.030940	-0.028604	0.163904	1.000000	
6 E.-m. sen. breadth	-0.005818	0.001348	0.029338	0.254373	-0.100985	1.000000

Second, Table 3.5 presents model-free evidence in the form of the raw correlation between the independent and dependent variables, including their respective p-values for statistical significance. The model-free evidence reveals significant and consistent positive correlations between some imitation and performance variables. The mode-free evidence furthermore shows a consistent and significant negative correlation between some early-mover and performance variables. Consequently, the model-free evidence provides initial support for our expected hypotheses by supporting a positive imitation-performance relationship and a negative early-mover-performance relationship in innovation contexts. We will refer back to the model-free evidence when validating the main results.

Table 3.5. Model-free Evidence over Various Time Periods

	Returns			**Sales Growth**			**EBIT Growth**		
Variable	1 Year	2 Year	5 Year	1 Year	2 Year	5 Year	1 Year	2 Year	5 Year
Im. mentions	0.04501***	-0.01431	-0.00159	-0.00250	-0.01431	-0.01358	0.00493	-0.00542	0.00705
Im. sentiment	-0.01601	0.02431*	-0.00396	0.02474*	0.02431**	-0.00245	0.03045**	0.00945	-0.02216
Im. sen. breadth	-0.04961***	0.00500	-0.00916	0.02150	0.00500	-0.04098***	-0.00539	-0.02018	-0.06054***
E.-m. mentions	-0.08014***	-0.02768**	-0.03114*	-0.03592***	-0.02768**	-0.01588	-0.04197***	-0.03160**	-0.05318***
E.-m. sentiment	-0.01879	0.00957	-0.10330***	0.01125	0.00957	0.01626	-0.00757	0.00370	0.02011
E.-m. sen. breadth	-0.00279	-0.00082	-0.04126**	0.00858	-0.00082	-0.03499**	-0.00943	0.00173	-0.03472**

Correlations are estimated with the Spearman's rank correlation coefficient. Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01).

3.5 Model Specification

Our models are estimated by robust pooled, fixed-effects and random-effects Ordinary Least Squares (OLS) panel regressions.[4] In the following subsections, we explain (1) how our observations and variables are combined into a panel data set, (2) our model selection process and how it is suited to analyse the imitation-performance relationship, and (3) describe the statistical process of identifying the appropriate panel regression model via outlier handling, test of the Multiple Linear Regression (MLR) assumptions and respective following decisions and statistical tests.

3.5.1 Panel Data

First, our panel data set is created by combining the empirical data with the performance outcome for every year and company. Here, our empirical data consists of all imitation and early-mover variables for each company and observation period combinations where valid annual reports are available. For every company and year, the annual, 2-year and 5-year financial market and operating performance outcomes are included upon sufficient data

[4]See OLS Regression Types in the Appendix for a brief comparison of the different OLS models.

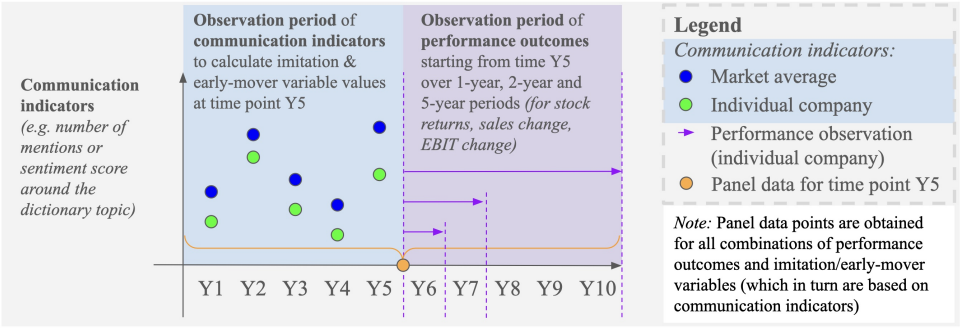

Figure 3.5. Panel Data Observation and Performance Periods

availability for the respective year and future period. Figure 3.5 describes how observation and outcome periods are combined in yearly data points. In summary, the panel data consists of a time series of imitation and early-mover observations, as well as performance outcomes for each entity. Table 3.6 shows an exemplary data set with the time series data from different entities. In our case, an entity refers to an individual company and is identified by its respective CIK.

Table 3.6. Sample Data Table

CIK	Year	Im. mentions	Im. sentiment	Im. sen. breadth	E.m. mentions	E.m. sentiment	E.m. sen. breadth	Performance
1800	2013	-0.2	0.5	-0.4	-2531.28	0.222	-0.00883	20.44427
	…	…	…	…	…	…	…	…
	2022	0.1	-0.9	-0.7	-1544.86	0.169	0.00829	-20.6644
6201	2013	0.2	-0.3	0.3	-3090.28	0.094	0.13469	1332.704
	…	…	…	…	…	…	…	…
	2022	-0.8	0.1	-0.2	9361.14	0.254	0.11355	-29.1759
…	…	…	…	…	…	…	…	…

3.5.2 Model Selection Approach

In the following subsection, we describe the general approach of our model selection process and how it is suited to analyse the imitation-performance relationship.

To create theoretical insights from panel data, a panel regression is a useful tool (Torres-Reyna, 2007) commonly employed in strategic management to analyse performance outcomes over text-based variables (Nauhaus et al., 2021; Nguyen and Agrawal, 2023). Therefore we argue that a panel regression is appropriate to investigate the imitation-performance relationship, as it fits both the empirical context and theoretical state of the research field (Edmondson and Mcmanus, 2007). A panel regression is furthermore suitable, as it runs over the time and entity dimension, allowing for a multi-dimensional analysis (Torres-Reyna, 2007). A panel regression can be approached from different angles,

depending on certain characteristics and assumptions within the underlying data (Meuleman et al., 2013; Torres-Reyna, 2007). Our models are estimated by an OLS regression which minimises the sum of the squared differences between the observed and predicted values of the dependent variable based on the multiple independent variables (Meuleman et al., 2013). As OLS panel regressions should fulfil certain statistical assumptions to deliver robust results. We follow the suggestions of Meuleman et al. (2013), who provide a methodological framework to select an appropriate model. They build on the widely accepted approach of Berry (1993), where a robust OLS regression needs to fulfil certain MLR assumptions. Meeting the MLR assumptions ensures that the conclusions drawn from the panel regression are both credible and generalisable (Torres-Reyna, 2007; Wolf and Best, 2013).

Furthermore, we aim to design our model selection to be as simple and consistent as possible. We always start our selection process from the standard pooled OLS regression based on the fact that we collected the empirical observations for all entities over the identical time period. In cases where we need to decide between fixed-effects or random-effects regression, we also follow the instructions of Meuleman et al. (2013) and Torres-Reyna (2007).

Pooled, fixed-effects, and random effects regression differ by how they account for potential statistical differences in the underlying panel data.[5]

While our empirical data is collected for every entity over an identical time period, thus providing preliminary arguments for a pooled regression, selecting the appropriate model is still important. It is reasonable to assume the existence of differences within our panel data, across the entities, over time or both. Our sample of S&P 500 companies consists of individual companies from various industries with inherently different characteristics such as business models, cultures and intra-organisational behaviours potentially not captured by our independent variables. Our independent variables only measure the outcome of the imitation decisions. From a longitudinal time-perspective, there furthermore might have been events that affect all entities in the same way, such as economic crises, wars, or policy changes. Thus, from a theoretical and empirical perspective, a robust model selection between the different pooled, fixed-effects and random-effects panel regression is essential (Torres-Reyna, 2007).

3.5.3 MLR Assumptions

In the following subsection, we describe the statistical process of identifying the appropriate panel regression model via outlier handling, test of the MLR assumptions, and respective following decisions and inherent statistical test.

[5]See OLS Regression Types in the Appendix for a brief comparison of the different OLS models.

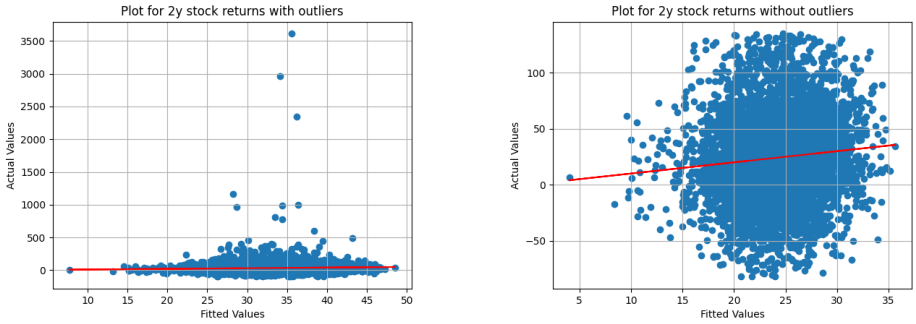

Figure 3.6. Comparison of Fitted Values and Actual Values for the Regression of 2-Year Stock Returns with Outliers (left) and Without Outliers (right).

We select the appropriate panel regression models by building on Berry (1993) and following Meuleman et al.'s (2013) suggestion of answering the question of how to handle the data-outliers and then consider a model as robust if it passes the five most important and widely cited MLR assumptions. We demonstrate the assumption testing process exemplary for the imitation-2-year-stock-return regression.

We start with our model selection process with a standard pooled OLS regression and answer the principal question if outliers should remain within the sample or be removed from it. Figure 3.6 compares the 2-year stock performance plot with and without outliers. For the right plot, outliers have been removed by applying the Interquartile Range (IQR) method. The IQR method is a robust and standard procedure in statistics to define the boundaries depending on the data distribution (Wilcox, 2012). We argue that the visual comparison suggests removing outliers, favouring a more homogenous distribution. Qualitatively, this decision is supported by our reasoning that companies who experience a 35-times multiplication (see Figure 3.6) of their stock value during a short period of time might not be in a normal phase of their life (e.g. major stock value increase in the year a company is listed for the first time). Consequently, we decide to remove the outlier data points. It is worthwhile to note that we do not remove the entire company from the sample, but only the respective data point as each entity might still provide observation periods with sufficient data points.

Based on the new outlier-free sample, all subsequent models are tested according to the most important MLR assumptions (Meuleman et al., 2013). In the following we test if the assumptions of (1) linearity, (2) normality, (3) multicollinearity, (4) independence of residuals and (5) homoscedasticity hold. We follow the approach of Meuleman et al. (2013), Berry (1993) and Torres-Reyna (2007) and test linearity with a standard plot, normality

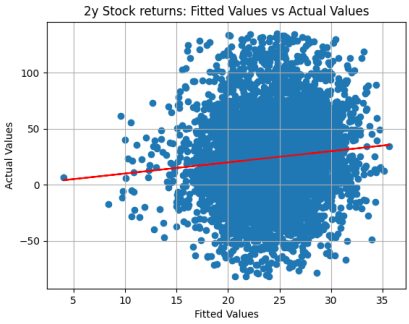

Figure 3.7. Linearity: Fitted Values vs Actual Values for the Regression of 2-Year Stock Returns in the LDA Innovation Context.

with the Quantile-Quantile (Q-Q) plot, multicollinearity with the Variance Inflation Factors (VIF)-test supported by the descriptive statistics of the variables, independence of residuals with the Durbin-Watson (DW) test and homoscedasticity with the Breusch-Pagan test. In cases of a failed Breusch-Pagan test, we use the Hausman test to estimate if a random-effects or fixed-effects model might be more appropriate than a pooled regression to account for the heteroscedasticity and subsequently result in a model selection that is able to pass this assumption.

We again demonstrate the assumption testing process exemplary for the imitation-2-year-stock-return regression.

First, a linear relationship between the dependent and independent variables is the key assumption for an OLS regression. We test this assumption using a standard plot as recommended by Meuleman et al. (2013), passing the test because of the linear trend depicted in Figure 3.7. We follow the argument of Itaoka (2012), and despite having a very low coefficient of determination (R^2) value and a resulting scattered plot, which is common when predicting stock returns, we do not identify a pattern suggesting deviations from a linear relationship. We argue the linearity assumption is met as our emphasis lies on testing hypotheses rather than predicting financial performance. Here, a low R^2 value is deemed appropriate for our primary objective of understanding relationships between variables. Additionally, statistically significant variables create causality around the phenomena under observation through their respective P-values and coefficients, maintaining robustness, as argued for in our thesis context.

Second, the multivariate normality of the regression residuals is assumed to ensure the validity of statistical inferences. We examine this assumption through a Q-Q plot following the approach of Meuleman et al. (2013). The residuals broadly follow the

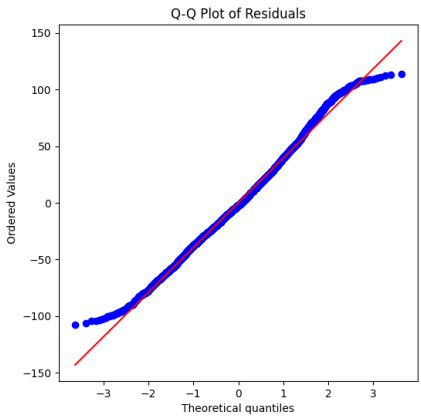

Figure 3.8. Normality: Q-Q Plot for the Regression Residuals of 2-Year Stock Returns in the LDA Innovation Context.

diagonal line in Figure 3.8. However, there are some slight deviations from the normal distribution in the tails.

Third, no multicollinearity between the independent variables is assumed to maintain the stability and efficiency of parameter estimates. We utilise the VIF-test as recommended by Meuleman et al. (2013) and pass because of the low VIF values observed.

Fourth, the independence of the observations should hold to prevent biased parameter estimates and inflated standard errors. We employ the Durbin-Watson (DW) test to assess this assumption, following the approach of Meuleman et al. (2013). We pass the test because the DW statistic falls within the acceptable range.

Fifth, homoscedasticity of the regression residuals is assumed for unbiased and efficient estimations. We test this assumption using the Breusch-Pagan test as per Meuleman et al. (2013). We pass the test because of the insignificant result, indicating homoscedasticity. In cases of observed heteroscedasticity, we would employ the Hausman test. The Hausman test evaluates the null hypothesis if the preferred model should be a random-effects regression or a fixed-effects regression.

Lastly, regarding the general robustness of the regression estimators, we follow common practices in management research. While we acknowledge that significance levels of models are arbitrary and context-dependent, we apply 10%, 5% and 1% significance levels as commonly used by well-reputed management journals in sentiment-based management contexts with financial outcome measures (Li et al., 2013; Nauhaus et al., 2021) as well

as innovation-centred dynamic capability research (Guerrero and Siegel, 2024). The described research contexts furthermore align with our research question of testing the dynamic capability-related imitation-performance relationship in innovation contexts.

3.6 Results

This section describes the significant results on the impact of imitation and early-mover behaviour on firms' financial performance outcomes — stock returns, sales change, and EBIT change— over three different time spans in innovation contexts using robust statistical pooled, fixed-effects, and random-effects OLS estimators where the model choice is supported by the MLR assumption and the Hausman test. The empirical study results are shown in Table 3.7.

We accept both hypothesis 1 and hypothesis 2: A higher degree of imitation behaviour has a positive effect on firm performance, while being an early-mover has a negative effect on firm performance.

Overall, the regressions estimate significant positive relationships between imitation behaviour and firm performance and significant negative relationships between early-mover behaviour and firm performance over various time periods. Although the R^2 values are generally small, all significant predictors consistently indicate the same type of relationship between the independent and dependent variables. Specifically, imitation mentions and imitation sentiment are linked to positive performance outcomes, while early-mover mentions, early-mover sentiment and early-mover sentiment breadth are linked to negative outcomes. The results are shown in Table 3.7. In Figure 4.1 of the analysis section, we provide the reader with a simplified summary of the nature of significant relationships.

Below, we describe the robust estimated performance outcomes of (1) imitation and early-mover mentions, (2) imitation and early-mover sentiment, and (3) imitation and early-mover sentiment breadth with their respective implications for the evaluation of the hypotheses as well as (4) the general regression outcomes. Significance levels are represented as $p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$, while (+) and (-) indicate a positive and negative relationship, respectively.

First, imitation mentions (Im. mentions) consistently significantly positively affect 1-year returns (+**) and 2-year returns (+**). Imitation mentions do not significantly affect sales change or EBIT change. Early-mover mentions (E.-m. mentions) consistently significantly negatively affect 1-year returns (-*), 2-year returns (-***), 1-year sales change (-***), 2-year sales change (-***), 1-year EBIT change (-***) and 2-year EBIT change (-**). The consistently significant positive effects of imitation mentions suggest the acceptance

of H1. The consistently significant negative effects of early-mover mentions suggest the acceptance of H2. The combination of consistently significant opposite outcome effects of imitation mentions and early-mover mentions further supports the acceptance of both H1 and H2.

Second, imitation sentiment (Im. sentiment) consistently significantly positively affects 5-year returns (+**) and 2-year sales change (+**). Imitation sentiment does not significantly affect EBIT change. Early-mover sentiment (E.-m. sentiment) consistently significantly negatively affects 2-year returns (-***), 5-year returns (-***), 2-year EBIT change (-***) and 5-year EBIT change (-***). Early-mover sentiment does not significantly affect sales change. The consistently significant positive effects of imitation sentiment suggest the acceptance of H1. The consistently significant negative effects of early-mover sentiment suggest the acceptance of H2. The combination of consistently significant opposite outcome effects of imitation mentions and early-mover mentions further supports the acceptance of both H1 and H2.

Third, imitation sentiment breadth (Im. sen. breadth) consistently significantly negatively affects 2-year returns (-***), 5-year sales change (-***) and 5-year EBIT change (-***). Early-mover sentiment breadth (E.-m. sen. breadth) significantly negatively affects 5-year sales change (-***). Early-mover sentiment breadth does not significantly affect Returns and EBIT change. While the imitation sentiment breadth variables themselves are consistently significant with the respective outcome, the communication indicator sentiment breadth itself requires further discussion, as described in the section 3.3.2. The identical significant negative outcome effects for both imitation and early-movers should be the focal point of the variable discussion.

Fourth, the regressions' overall respective adjusted R^2 values indicate that, in most cases, the estimators explain a small but significant variance in the variables across different time frames and models. The number of observations varies across the models depending on the outcome variable, observation window, data availability and outlier removal. P-values for the models differ, where most models are significant, with several models achieving high levels of significance ($p < 0.01$) and only the 1-year return regression being overall not significant ($p=0.1289$). All results are estimated with robust standard errors and are estimated as described in the methodology section of the thesis.

In summary, when significant, the results show a consistent positive effect of imitation mentions and sentiment on performance outcomes and a consistent negative effect of early-mover mentions and sentiment on performance outcomes over time. While the sentiment breadth variable leaves room for further discussion, we demonstrate a consistent relationship between the key imitation and early-mover variables. Consequently, we accept hypothesis 1, that a higher degree of imitation has a positive effect on firm

performance, and hypothesis 2, that being an early-mover has a negative effect on firm performance.

Table 3.7. Results over Various Time Periods for LDA Innovation

	Returns			Sales Change			EBIT Change		
Variable	1 Year	2 Year	5 Year	1 Year	2 Year	5 Year	1 Year	2 Year	5 Year
const	13.635***	24.968***	79.230***	0.0583***	0.1107***	0.2957***	0.0544***	0.1229***	0.3469***
	(0.3783)	(0.6034)	(2.9850)	(0.0015)	(0.0028)	(0.0068)	(0.0040)	(0.0044)	(0.0084)
Im. mentions	1.7476**	2.3889**	3.7007	0.0005	0.0020	-0.0055	0.0042	0.0028	0.0097
	(0.8402)	(1.1484)	(2.6155)	(0.0029)	(0.0053)	(0.0128)	(0.0057)	(0.0099)	(0.0185)
Im. sentiment	0.3753	0.0412	5.4140**	0.0043	0.0123**	-0.0004	0.0086	0.0195**	-0.0282
	(0.8066)	(1.1000)	(2.4218)	(0.0028)	(0.0051)	(0.0123)	(0.0054)	(0.0094)	(0.0182)
Im. sen. breadth	-0.0946	-3.9084***	-1.0114	0.0013	0.0022	-0.0373***	0.0034	-0.0016	-0.0463***
	(0.7418)	(1.0332)	(2.2247)	(0.0026)	(0.0048)	(0.0115)	(0.0050)	(0.0087)	(0.0165)
E.-m. mentions	-0.0003*	-0.0004***	-0.0001	-1.047e-06***	-2.01e-06***	-1.926e-06	-2.34e-06***	-4.224e-06**	-4.32e-06
	(0.0002)	(0.0001)	(0.0003)	(2.832e-07)	(5.192e-07)	(1.241e-06)	(6.597e-07)	(6.597e-07)	(3.829e-06)
E.-m. sentiment	-5.0237	-6.6681***	-51.638***	0.0101	0.0153	0.0139	0.0164	-0.1448***	-0.1551*
	(3.8783)	(2.5299)	(9.1117)	(0.0064)	(0.0116)	(0.0274)	(0.0157)	(0.0461)	(0.0895)
E.-m. sen. breadth	5.9852	-0.6104	-26.228	0.0075	0.0134	-0.1608***	0.0153	0.1156	0.0531
	(8.1473)	(5.0661)	(18.878)	(0.0127)	(0.0232)	(0.0553)	(0.0312)	(0.0954)	(0.1879)
R^2 adj.	0.0022	0.0092	0.0252	0.0036	0.0045	0.0060	0.0026	0.0052	0.0044
N° observations	5009	4895	3198	4719	4811	4390	4404	4515	4167
P-value	0.1289	0.0000	0.0000	0.0087	0.0015	0.0002	0.0711	0.0025	0.0153
Regression type	Fixed-effects entity & time	Pooled	Random-effects	Pooled	Pooled	Pooled	Random-effects	Fixed-effects entity & time	Fixed-effects entity & time

Models are estimated using pooled OLS, fixed-effects OLS, or random-effects OLS estimators with robust standard errors (in parentheses). The model selection is based on the MLR assumption and Hausman test, as specified in Section 3.5.4. Variables are abbreviated as "Im." for Imitation and "E.-m." for Early-mover, with respective mentions, sentiment (sen), and sentiment breadth. Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01).

3.7 Quality Discussion

Based on the mature state of the RBV, dynamic capability and economic herding literature, as well as the theoretical tensions on the expected outcomes of imitation performance-relationship, we argue for the need to critically reflect on the quality of the empirical study and inherent operational choices (Edmondson and Mcmanus, 2007) to better assess the potential theoretical contributions of our results (Daft, 1983).

Concretely, we reflect on (1) the reliability, (2) the validity from a measurement, internal, external and ecological perspective and lastly (3) the replicability of our study.

First, reliability ensures that the measures used in the study are stable and consistent over time (Saunders et al., 2009). We argue that our results are reliable as all individual variables are consistent, when significant, over all time periods and all outcome measures. From an internal reliability point of view, we argue that our multiple variables within the individual concepts are also reliable, as most of the robust measures of the mentions and sentiment consistently align with opposite outcomes. By building the study on a positivist approach while relying on quantitative empirical data, we argue our results are largely not affected by inter-observer consistency risks (Bell et al., 2022; Saunders et al., 2009; Straub and Gefen, 2004). The only cause of inter-observer risk might be created when selecting the dictionaries. Furthermore, our very weakly correlated independent variables (see Table 3.4) and robust model selection based on the most common MLR assumptions allow us to expect reliable results.

Second, validity refers to the integrity of the results (Creswell and Creswell, 2017). We evaluate the validity of our study from a measurement, internal, external and ecological perspective.

Measurement validity ensures that the measures used in the study reflect the researched concepts (Creswell and Creswell, 2017). While our concrete mathematical definitions have mostly not been used in the literature before, primarily due to the lack of empirical studies on the imitation-performance relationship (Posen et al., 2023), we developed our variables with the goal of capturing the definition of imitation as accurately as possible. We base the imitation and early-mover measurement on the established strategy, imitation and herding literature and thus argue for measurement validity. To mitigate the risk of novel variables and cater to the mature research state, we furthermore ground the evaluation of the hypotheses in the most robust mentions and sentiment variable while reserving the most abstract variable of sentiment breadth for extended discussion. Lastly, the outcome measures are objective and diverse, accurately measuring the performance of the observed firms.

Internal validity examines the causal relationships between variables, ensuring that the independent variable influences the dependent variable (Saunders et al., 2009). We argue for sufficient internal validity from being outcome-oriented as the independent variables are backwards-looking and the dependent ones are forward-looking. Only then do we test for causal relationships with robust statistical models. The panel regression models not only provide consistent results over the various outcome measures and time periods when significant but are also consistent with the significant correlations from the model-free evidence, as shown in Table 3.5. We aimed to validate the study by evaluating the imitation-performances relationship in varying contexts. Additionally, in the LDA innovation dictionary, we run robust stock return panel regression for the DIS innovation context, the LDA uncertainty context, and the LDA sustainability context. Yet, as Table 3.8 reveals, the DIS innovation context does not necessarily validate our study, being largely insignificant. Only the early-mover mentions and early-mover sentiment are significant. While the early-mover sentiment aligns with the main results, the early-mover mentions contradict them. However, as predicted by Miric et al. (2023), existing dictionaries might not be able to capture trends accurately and thus fail to provide a robust empirical context. The regressions on the LDA uncertainty and LDA sustainability reveal interesting insights. We return to those dictionaries in the post-hoc section of this thesis. In summary, we argue for sufficient internal validity based on our model-free evidence in combination with robust MLR-assessed panel regressions.

External validity evaluates if results can be generalised beyond the observed sample size. We argue for external validity as our S&P 500 company sample is already diverse across industries and sectors, with revenues generated globally, reducing the risk of firm-specific effects. Additionally, we also observe companies over the entire sample period and continuously track companies that have been added to or removed from the S&P 500 to mitigate sample-specific survivorship bias. Furthermore, we try to mitigate remaining firm-specific effects by using fixed-entity effects panel regression models when suggested by the MLR assumptions. Combined with diverse and objective performance measures, we not only avoid a common pitfall in management research when trying to resolve theoretical frictions between mature literature streams (Miller and McKee, 2021; Samba et al., 2021) but further promote the external validity and subsequent potential generalisability of our results. In summary, we argue for high generalisability when theorising about large corporations in the Western economy.

Ecological validity assesses whether the results are transferable from theory to practice. While all our empirical data measures the behaviour and outcome of the sampled firms, we argue for reduced ecological validity value, as our adjusted R^2 regression values are consistently low. Our study results aim to contribute to the academic dialogue rather than to inform concrete managerial decisions. Our aggregated performance lens does not

Table 3.8. Results over Various Time Periods for DIS Innovation

	Returns		
Variable	1 Year	2 Year	5 Year
const	13.559***	24.015***	79.465***
	(0.4450)	(0.8829)	(2.9899)
Im. mentions	-0.0275	0.1027	-2.6175
	(0.7048)	(1.0492)	(2.0841)
Im. sentiment	-0.0563	-0.2542	0.1108
	(0.7354)	(1.0703)	(2.1675)
Im. sen. breadth	-0.2015	-1.3275	-0.1409
	(0.6764)	(0.9782)	(2.0873)
E.-m. mentions	0.0003	-4.811e-05	0.0077**
	(0.0005)	(0.0010)	(0.0032)
E.-m. sentiment	-0.8778	-2.7014	-9.4839*
	(0.9068)	(1.6882)	(4.9321)
E.-m. sen. breadth	-2.4916*	-1.7292	-10.041
	(1.4797)	(2.8000)	(7.9809)
R^2 adj.	0.0008	0.0003	0.0160
N° observations	5009	4895	3198
P-value	0.6856	0.9742	0.0000
Regression type	Random-effects	Random-effects	Random-effects

Models are estimated using pooled OLS, fixed-effects OLS, or random-effects OLS estimators with robust standard errors (in parentheses). The model selection is based on the MLR assumption and Hausman test, as specified in Section 3.5.4. Variables are abbreviated as "Im." for Imitation and "E.-m." for Early-mover, with respective mentions, sentiment (sen), and sentiment breadth. Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01).

allow us to derive concrete managerial conclusions for individual firms on how to behave in a specific situation.

Third, replicability refers to the reproducibility of the results (Bell et al., 2022). We argue for high replicability, as our study is based on a positivist, empirical, and statistically sound approach. We furthermore provide the reader with a detailed description of our operational choices. Additionally, we provide all LDA dictionaries around innovation, uncertainty and sustainability for all observed years from 2009-2022 in the Appendix . Based on our panel regressions and previous conclusions on sufficient reliability and validity, we expect similar methods and measurements to replicate similar findings.

In summary, we made the operational choices with the ambition of creating reliable, valid and replicable results. While our research approach involves techniques such as LDA topic modelling and sentiment analysis on text data, we aim to ensure validity

with robust models and model-free evidence. We furthermore design our measures to capture the imitation-performance relationship accurately. We consider our sample size comprehensive, and thus, our results are relatively generalisable. Lastly, we deem our study replicable because it is rooted in a transparent and positivist approach, supporting the potential likelihood of reproducing findings.

4

Analysis & Discussion

In the following section, we discuss the empirical findings that a higher degree of imitation behaviour has a positive effect on firm performance while being an early-mover has a negative effect on firm performance within the context of innovation. We analyse the results through (1) our conceptual framework and inherent theories, (2) elaborate on the potential explanations between the aligned and contradicting empirical variables, and (3) post-hoc contribute with an analysis of the outcomes of imitation and early-mover behaviour in the context of uncertainty and sustainability.

4.1 Imitation-Performance Relationship

Strategic imitation is the "attempt to reproduce [...] another firm's product, processes, capabilities, technologies or [strategic] decision [..]" (Posen et al., 2023, p 76). While the imitation-performance relationship was empirically unexplored and lacking a clear understanding of the expected diverse performance outcomes, management research further developed into two very different theoretical schools of thought, on the one hand implying a static understanding with adverse performance effects and on the other hand a dynamic understanding with expected positive outcomes (Posen et al., 2023). The static model aligns with the RBV and the dynamic model with the dynamic capability literature (Barney, 1991; Posen et al., 2023; Teece et al., 1997). As shown by this study, imitation furthermore overlaps with economic herding theories (Banerjee, 1992; Bikhchandani et al., 1992) and can be linked to the early-mover discussion (Hoppe, 2000; Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005).

Overcoming the theoretical tensions between the static and dynamic imitation models, this study reveals that the imitation-performance relationship should be approached from a firm-centric aggregated lens to investigate whether imitation leads to negative or positive performance outcomes. When measuring imitation as the observable correlated

behaviour of a group of organisations, occurring consciously or unconsciously, as they pursue similar opportunities or evaluate topics and trends in similar manners at the same time against the diverse performance outcomes of shareholder value creation, revenue growth and profitability growth our empirical study reveals:

A higher degree of imitation behaviour has a positive effect on firm performance while being an early-mover has a negative effect on firm performance.

In the following, we analyse the (1) positive imitation-performance and (2) negative early-mover relationship as suggested by our conceptual framework by applying the dynamic capability literature, RBV and economic herding theories.

First, a positive imitation-performance relationship can be explained through a dynamic capability perspective, where dynamic capabilities include effectively sensing, seizing, and transforming resources (Teece et al., 1997). Consequently, our results support viewing imitation not merely as a reactive strategy but as a key enabler of dynamic capabilities enhancing a firm's adaptive capabilities, as suggested by Posen et al. (2023). In this context, strategic imitation can serve as a mechanism for firms to acquire and integrate external innovations that complement their existing capabilities, thereby fostering an environment of continuous learning and adaptation on an aggregated market level (Cohen and Levinthal, 1990; Levitt and March, 1988). By purposefully imitating, firms do not just copy but intelligently integrate these practices with their unique capabilities, leading to innovative combinations (Posen et al., 2023). This integration could, for example, enhance their responsiveness to market changes and technological advancements, leading to improved performance outcomes (Eisenhardt and Martin, 2000; Posen et al., 2023). Furthermore, our findings and the theoretical conclusion to consider imitation as a dynamic and intentional imitation process and not just as a dyadic interaction between imitator and target but as a broader market-level process strengthens the theoretical relationship between the imitation literature and adjacent management constructs: Firms can discern patterns, absorb knowledge, and diffusing trends that signal successful strategies (Cohen and Levinthal, 1990; DiMaggio, Powell, et al., 1983; Jaffe, Trajtenberg, and Fogarty, 2000; Levitt and March, 1988; Strang and Soule, 1998).

Integrating economic herding theories offers further insights into why imitation could positively influence firm performance. Economic herding, often seen negatively due to its association with irrational conformity, can also be interpreted as a rational response under conditions of uncertainty and incomplete information (Banerjee, 1992; Bikhchandani et al., 1992). When firms observe a successful strategy or innovation adopted by peers, imitating such behaviours can be a strategic decision rather than a blind follow-the-crowd mentality. Thus, strategic imitation would allow firms to minimise risks and costs associated with developing untested strategies while capitalising on the proven

success of others. In summary, economic herding theory suggests that strategic imitators achieve better outcomes as they combine private and public information rather than disregarding their private information.

Consequently, in the dynamic model and according to herding theories, a firm would not converge to the individual performance maximum, but the suboptimal equilibrium that might still be more beneficial than relying solely on private information. This view aligns with Posen et al.'s (2023) dynamic imitation model, where the imitator integrates its own and competitors' knowledge, processes, and technologies. While investor herding in financial markets can still lead to a negative performance (Jiang and Verardo, 2013), we do not see this as contradictory to our results of a positive imitation-performance relationship. We base this conclusion on the assumption that firm performance is not strictly tied to financial market liquidity and, thus, prevalent differences in financial markets and strategic decision-making contexts when aiming for competitive advantages.

Second, the negative early-mover performance relationship can be explained by applying a dynamic capability perspective supported by our findings on the imitation-performance relationship. Early-mover effects describe both the strategic advantages and costs of being first or among the first when e.g. entering a market or adopting a new technology (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). Our results are partly contradicting the traditional RBV (Barney, 1991), as negative performance outcomes might be rooted in the fact that early-movers may initially seize substantial market share and secure key resources but also face the burden of uncertainty and the high costs of pioneering new markets or technologies (Lieberman and Montgomery, 1988). Early-movers often bear the brunt of educating the market, refining immature technologies, and overcoming initial market resistance—costs that later entrants can avoid (Lieberman and Montgomery, 1988). This scenario suggests that while early-movers initially establish a competitive stance, the costs and risks of these pioneering moves can erode their advantages. Simultaneously, imitators can learn from the early-mover's experiences and enter the market with fewer uncertainties and lower costs. Later entrants and imitators can leverage early-movers' trials and errors, incorporating external knowledge to sidestep pitfalls and enhance their strategic positioning (Eisenhardt and Martin, 2000; Suarez and Lanzolla, 2005). This conclusion is consistent with adjacent constructs, such as knowledge spillovers and the diminishing benefits that early-movers might accrue (Jaffe, Trajtenberg, and Henderson, 1993; Jaffe, Trajtenberg, and Fogarty, 2000). As early-movers innovate and break new ground, they inadvertently share valuable insights through various channels—industry reports, patents, executive speeches, and more—which later entrants can exploit without incurring the initial costs of innovation (Hoppe, 2000; Jaffe, Trajtenberg, and Henderson, 1993). The aggregated negative effect of being an early-mover is aligned with the dynamic understanding of imitation; firms do not just copy

but intelligently integrate competitors' practices with their unique capabilities, leading to innovative combinations (Posen et al., 2023) without bearing the full cost (Levitt and March, 1988; Terlaak and Gong, 2008).

In conclusion, a positive relationship between imitation and performance is supported by integrating dynamic capabilities that enhance adaptive strategies and economic reasoning, assuming improved search outcomes through an intentional approach to leverage industry-wide insights for a strategic advantage.

4.2 Sentiment Deviation

In this subsection, we extend the discussion on the significant independent variables by dissecting the (1) mentions and sentiment variables from the (2) sentiment-breadth variable. To do so, we simplify and visualise the empirical results of the imitation-performance relationship and early-mover-performance relationship in Figure 4.1.

First, the significant imitation mentions and imitation sentiment variables have a consistent positive impact on all performance variables over all time periods. The significant early-mover mentions and early-mover sentiment variables have a consistently negative impact on all performance variables over all time periods. These consistent opposite outcomes for imitators and early-movers via the mentions and sentiment variables strongly align with the previous theoretical analysis. These results support that imitation indeed is a key driver of innovation and results in a sustained competitive advantage (Posen et al., 2023), while early-movers subsequently experience adverse performance outcomes.

		Returns			Sales change			EBIT change		
		1Y	2Y	5Y	1Y	2Y	5Y	1Y	2Y	5Y
Imitation	Mentions	**	**							
	Sentiment			**		**			**	
	Sentiment breadth		***				***			***
Early-mover	Mentions	*	***		***	***		***	**	
	Sentiment		***	***					***	*
	Sentiment breadth						***			

Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01). The green and red colours, respectively, indicate a positive and negative relationship. The complete and robust regression results can be found in Table 3.7

Figure 4.1. Simplified Results over Various Time Periods for LDA Innovation

Second, the sentiment breadth variable results deviate from the previous mentions and sentiment variable results. While the previous significantly consistent relationship and expected opposite outcomes between imitator and early-movers help to advance the theoretical understanding of imitation, our study also shows that the standard deviation with the sentiment (sentiment breadth variables) consistently predicts a negative effect on performance across all variables. This finding is unexpected, as both imitators and early movers experience identical outcomes. This revelation calls for further discussion on why both imitators and early-movers experience worse performance outcomes when having higher deviation in their sentiment around innovation topics. A simple explanation might be the novel nature of the sentiment breadth variable. With sentiment breadth not used previously in the mature state of the literature before and the sentiment breadth variable being abstract in itself, it is possible that the variable fails to capture the phenomena under observation (Edmondson and Mcmanus, 2007).

In summary, the literature-grounded mentions and sentiment variable results support the theoretical claims, while the novel sentiment breadth variable paths the way for further theorising beyond the scope of this thesis.

4.3 Post-Hoc

In this section, we post-hoc analyse the imitation and early-mover performance relationship within uncertainty and sustainability contexts. The post-hoc contexts are defined by the respective LDA dictionaries. The findings reveal a surprising divergence: While performance outcomes under uncertainty largely align with the prior innovation contexts, the sustainability context unexpectedly indicates a positive performance impact for early-movers, especially those who address negative sentiment topics. These results suggest that addressing sustainability topics may provide distinct advantages, contrasting with the typically negative early-mover effects observed in innovation and uncertainty contexts. Concretely, in this section we (1) describe the post-hoc results on the imitation and early-mover performance relationship for the uncertainty and sustainable contexts and (2) discuss the respective findings with the primary focus on the sustainability context.

First, the post-hoc performance outcomes of the variables used within the panel regression differ between the LDA uncertainty and sustainability dictionaries.

In the uncertainty context (see Table 4.1), results - when significant - mostly align with previous findings. Imitation mentions (Im. mentions) and imitation sentiment (Im. sentiment) are statistically insignificant. Imitation sentiment breadth significantly negatively affects 2-year returns (-*). Early-mover mentions (E.-m. mentions) and early-mover sentiment (E.-m. sentiment) positively affect 1-year returns (+) and 5-year returns

(+***). Early-mover sentiment breadth (E.-m. sen. breadth) positively affects 2-year returns (+*). Overall, the 1-year and 2-year returns regressions exhibit a low adjusted R^2 value with a correspondingly high p-value. However, for the 5-year returns, the adjusted R^2 value is significantly higher, with a very low P-value (p<0.01), which suggests that the estimators explain a small but significant variance in the performance outcome.

Table 4.1. Results over Various Time Periods for LDA Uncertainty

	Returns		
Variable	1 Year	2 Year	5 Year
const	13.866***	24.554***	78.609***
	(0.3815)	(0.5379)	(2.9633)
Im. mentions	-0.7967	-1.3832	-3.5581
	(0.8667)	(1.2221)	(2.5809)
Im. sentiment	-0.8261	0.2978	2.2235
	(0.8517)	(1.2001)	(2.6684)
Im. sen. breadth	-1.2114	-2.0099*	1.8205
	(0.7856)	(1.1086)	(2.4201)
E.-m. mentions	-0.0003*	-0.0001	-0.0002
	(0.0002)	(0.0003)	(0.0004)
E.-m. sentiment	-0.8120	-9.5369	-28.098***
	(4.7041)	(6.6440)	(10.687)
E.-m. sen. breadth	9.0686	23.928*	-9.8812
	(9.6052)	(13.607)	(22.712)
R^2 adj.	0.0018	0.0021	0.0159
N° observations	5009	4895	3198
P-value	0.2544	0.1713	0.0000
Regression type	Fixed-effects entity & time	Fixed-effects entity & time	Random-effects

Models are estimated using pooled OLS, fixed-effects OLS, or random-effects OLS estimators with robust standard errors (in parentheses). The model selection is based on the MLR assumption and Hausman test, as specified in Section 3.5.4. Variables are abbreviated as "Im." for Imitation and "E.-m." for Early-mover, with respective mentions, sentiment (sen), and sentiment breadth. Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01).

In the sustainability context (see Table 4.2), results differ partially from the previous outcomes. This is surprising as only the context changes from innovation (and uncertainty) to sustainability, while the method remains identical. We find that imitation mentions (Im. mentions) still positively affect 5-year returns (+**). Imitation sentiment (Im. sentiment) still positively affects 2-year returns (+*) and 5-year returns (+***). Imitation sentiment breadth is statistically insignificant for all three performance outcomes. New is that early-mover mentions (E.-m. mentions) positively affect 2-year returns (+*)

and 5-year returns (+***). Early-mover sentiment (E.-m. sentiment) negatively affects 2-year returns (-*) and 5-year returns (-**). Early-mover sentiment breadth (E.-m. sen. breadth) is statistically insignificant for all three performance outcomes. Similarly to LDA uncertainty, the 1-year and 2-year returns regressions exhibit a low adjusted R^2 value with a correspondingly high p-value. However, for the 5-year returns, the adjusted R^2 value is significantly higher, with a very low P-value ($p<0.01$), which suggests that the estimators explain a small but significant variance in the performance outcomes.

Table 4.2. Results over Various Time Periods for LDA Sustainability

	Returns		
Variable	1 Year	2 Year	5 Year
const	13.603***	24.520***	79.877***
	(0.4614)	(0.6428)	(1.4816)
Im. mentions	-0.0413	1.5872	5.0265**
	(0.8174)	(1.1542)	(2.2833)
Im. sentiment	1.0705	2.1113*	7.0904***
	(0.8191)	(1.1574)	(2.4177)
Im. sen. breadth	0.9016	-0.6300	1.7142
	(0.8770)	(1.2322)	(2.5887)
E.-m. mentions	1.123e-05	0.0001*	0.0006***
	(5.322e-05)	(7.449e-05)	(0.0002)
E.-m. sentiment	-0.1292	-10.103*	-24.524**
	(3.8968)	(5.4801)	(11.305)
E.-m. sen. breadth	-9.0845	6.2083	12.647
	(7.3652)	(10.362)	(20.609)
R^2 adj.	0.0011	0.0021	0.0108
N° observations	5009	4895	3198
P-value	0.5535	0.1670	0.0001
Regression type	Fixed-effects entity & time	Fixed-effects entity & time	Fixed-effects entity & time

Models are estimated using pooled OLS, fixed-effects OLS, or random-effects OLS estimators with robust standard errors (in parentheses). The model selection is based on the MLR assumptions and the Hausman test, as specified in Section 3.5.4. Variables are abbreviated as "Im." for Imitation and "E.-m." for Early-mover, with respective mentions, sentiment (sen), and sentiment breadth. Stars indicate statistical significance (*p < 0.1; **p < 0.05; ***p < 0.01).

Second, analysing the statistically significant observed results from both the LDA uncertainty and LDA sustainability dictionary, we find that, on the one hand, the uncertainty dictionary results broadly align with the findings from the innovation dictionary. On the other hand, the sustainability dictionary surprisingly estimates a positive performance relationship for early-movers who address sustainability topics and those who

address sustainability topics with negative sentiment.

For the uncertainty context, significant early-mover mentions and sentiment variables negatively affect performance, thus aligning with the innovation dictionary and potential underlying theoretical reasons on the adverse effects of being an early-mover (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). Contradicting the innovation dictionary, the imitation sentiment breadth and early-mover sentiment breadth estimate the opposite outcome, with imitation sentiment breadth negatively affecting and early-mover sentiment breadth positively affecting performance. The applicability of the novel sentiment breadth variable has, however, been discussed in section 3.3.2 and further elaborated in section 4.2, resulting in the suggested dedicated theoretical discussion beyond the scope of this thesis.

Analysing the LDA sustainability results and reflecting on the fact that the imitation variables estimate a significantly positive effect on performance while the early-mover sentiment variable consistently significantly negatively affects both 2-year and 5-year returns, we can conclude that the sentiment of sentences containing the words in the sustainability dictionary affect performance. We show that a firm addressing sustainability topics earlier realises improved performance outcomes. We also show that firms that express increased negative sentiment in sentences addressing sustainability experience a positive effect on performance. This finding deviates from previous results; the early-mover mentions variables positively affect 2-year returns and 5-year returns. This result indicates that addressing sustainability more than the market affects performance positively. We believe this post-hoc result is exciting as early-movers do not experience the adverse performance effect early-movers face in innovation contexts.

5

Conclusions

In the following section, we conclude our study while the subsections summarise our (1) theoretical contributions, (2) managerial implications, (3) research limitations, and (4) potential future research opportunities.

We conclude our study with the answer to the research questions on the robust and diverse performance outcomes of strategic imitation: A higher degree of imitation behaviour has a positive effect on firm performance while being an early-mover has a negative effect on firm performance within the context of innovation. We systematically analysed the imitation-performance relation for S&P 500 companies within the context of innovation during 2009-2023. To overcome the combined nascent friction within the static and dynamic understanding of imitation and a lack of empirical evidence on the imitation-performance relationship (Posen et al., 2023) we developed hypotheses based on the RBV, dynamic capability, and economic herding literature (Banerjee, 1992; Barney, 1991; Bikhchandani et al., 1992; Peteraf, 1993; Teece et al., 1997) to empirically explore (Bell et al., 2022; Edmondson and Mcmanus, 2007) the imitation-performance relationship. As previous empirical evidence is both scarce and contradicting (Banker et al., 2013; Barreto and Baden-Fuller, 2006; Deephouse, 1999; Giachetti, Lampel, et al., 2017; Giachetti and Pira, 2022; Posen et al., 2023), we utilise large-scale machine-learning enabled LDA and sentiment-based methods (Bayer et al., 2017; Bellstam et al., 2021; Blei et al., 2003; Li et al., 2013; Ligorio et al., 2022; Miric et al., 2023) in combination with objective and diverse performance measures such as shareholder value creation, revenue growth and profitability improvements (Berk and DeMarzo, 2007; Eberhart et al., 2004). We use statistically robust pooled, fixed-effects and random-effects panel regressions (Berry, 1993; Torres-Reyna, 2007; Wolf and Best, 2013). Our approach aims to overcome frictions between established management theories (Miller and McKee, 2021; Samba et al., 2021) and estimates a positive imitation-performance and negative early-mover-performance relationship in innovation context. We demonstrate post-hoc how, in contradiction to

innovation contexts, companies are financially rewarded for addressing sustainability topics early and in negatively connotated contexts.

5.1 Theoretical Contributions

We contribute to the literature by (1) answering the research questions that imitation positively affects diverse firm performance outcomes, overcoming frictions between the static and dynamic model of imitation, extending the discussion to early-mover performance relationships, and fostering a dialogue between management and economics research. (2) We overcome the lack of robust and diverse empirical evidence on the performance outcomes of imitation. (3) We methodologically contribute to management research by developing partly novel imitation measures while employing machine-learning-enabled sentiment analysis and LDA topic modelling on 10-K annual reports. (4) We provide the management field with topic dictionaries covering the 100 statistically most important words companies use when speaking about innovation, uncertainty, and sustainability for each year from 2009-2022. (5) We post-hoc demonstrate how being an early-mover in a sustainability context yields diverse performance improvements.

First, our results show that a higher degree of imitation behaviour has a positive effect on firm performance, while being an early mover has a negative effect on firm performance. Strategic imitation is the "attempt to reproduce [...] another firm's product, processes, capabilities, technologies, or [strategic] decision [...]" (Posen et al., 2023, p 76). While the imitation-performance relationship was empirically unexplored and lacked a clear understanding of the expected diverse performance outcomes, management research further developed into two very different theoretical schools of thought, on the one hand implying a static understanding with adverse performance effects and on the other hand a dynamic understanding with expected positive outcomes (Posen et al., 2023). The static model aligns with the RBV and the dynamic model with the dynamic capability literature (Barney, 1991; Posen et al., 2023; Teece et al., 1997). As shown by this study, imitation furthermore overlaps with economic herding theories (Banerjee, 1992; Bikhchandani et al., 1992) and can be linked to the early-mover discussion (Hoppe, 2000; Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005).

The observed positive imitation-performance relationship may be explained by applying a dynamic capability perspective (Teece et al., 1997). Our results support viewing imitation not merely as a reactive strategy but as a key enabler of dynamic capabilities (Posen et al., 2023), serving as a mechanism for firms to acquire and integrate external innovations that complement their existing capabilities. This process fosters an environment of continuous learning and adaptation at an aggregated market level (Cohen and Levinthal,

1990; Levitt and March, 1988), consequently improving performance outcomes by discerning patterns and absorbing knowledge (Cohen and Levinthal, 1990; DiMaggio, Powell, et al., 1983; Jaffe, Trajtenberg, and Fogarty, 2000; Levitt and March, 1988; Strang and Soule, 1998). A positive imitation-performance relationship furthermore helps to align Posen et al.'s (2023) dynamic model of imitation with economic herding theories, where the imitator integrates its own and competitors' knowledge, processes, and technologies and thus avoids potential adverse effects of disregarding private information (Banerjee, 1992; Bikhchandani et al., 1992). The positive imitation-performance relationship seeks to be a starting point for a variety of research questions within imitation research (Posen et al., 2023). The negative early-mover performance relationship may also be explained by applying a dynamic capability perspective supported by our findings on the nature of the imitation-performance relationship. Early-mover effects describe both the strategic advantages and costs of being the first or among the first when, e.g., entering a market or adopting a new technology (Lieberman and Montgomery, 1988; Suarez and Lanzolla, 2005). Our results are partly contradicting the traditional RBV (Barney, 1991), as negative performance outcomes might be rooted in the fact that early-movers may initially seize substantial market share and secure key resources but also face the burden of uncertainty and the high costs of pioneering new markets or technologies (Lieberman and Montgomery, 1988). Our findings support the argument that early-movers may initially establish a competitive stance, yet imitators can learn from their errors to enhance their strategic positioning (Eisenhardt and Martin, 2000; Suarez and Lanzolla, 2005) without bearing the early-mover costs of innovation (Hoppe, 2000; Jaffe, Trajtenberg, and Henderson, 1993).

Second, we empirically overcame the lack of robust and diverse evidence on the performance outcomes of imitation and bridged tensions between the static and dynamic imitation models hindering effective measurement (Miller and McKee, 2021; Posen et al., 2023; Samba et al., 2021). We showed that approaching the imitation-performance relationship from a firm-centric aggregated performance perspective, measuring imitation as observable correlated behaviour of a group of organisations (Hwang and Salmon, 2004; Nofsinger and Sias, 1999), combined with statistically robust panel regression models (Berry, 1993; Torres-Reyna, 2007; Wolf and Best, 2013) foster academic discourse. We filled the research gap by accounting for diverse performance outcomes, including shareholder value creation, revenue growth, and profitability growth for S&P 500 companies from 2009-2022 (Posen et al., 2023).

Third, we contribute methodologically to management research by developing partly novel communication indicators to measure imitation behaviour while using large-scale machine learning-enabled LDA and sentiment analysis to extract information value from textual data (Bayer et al., 2017; Bellstam et al., 2021; Blei et al., 2003; Li et al., 2013; Ligorio

et al., 2022; Miric et al., 2023). Our study furthermore serves as an example that pre-existing dictionaries potentially cause issues when transferred to other documents or time periods (Miric et al., 2023). Compared to the LDA-created innovation dictionary, the DIS innovation dictionary struggled to produce significant results.

Fourth, we provide researchers with our dictionaries created via LDA topic modelling. The dictionaries in Appendix A contain the 100 most important words S&P 500 companies use when speaking about innovation, uncertainty, and sustainability. All dictionaries cover each year for the period 2009-2022.

Fifth, we post-hoc demonstrated how being an early-mover in a sustainability context yields diverse performance improvements. Furthermore, firms that address sustainability in negatively connotated contexts early also experience positive performance outcomes.

In summary, we contribute to the academic discussion by answering the research question and providing evidence that imitation behaviour positively affects diverse performance outcomes. While centred around the theoretical questions of the imitation-performance relationship, we believe some of the conducted work, especially the created dictionaries, can contribute to management research as a whole.

5.2 Managerial Implications

Reflecting on the managerial implications and potential value for practitioners, we diagnose limited direct applicability of our results in specific decision contexts due to the (1) statistical characteristics of our study (low adj. R^2) and (2) the nature of our research question, yet we (3) provide the reader with a high-level strategy rule to guide managerial thinking on strategic imitation and sustainability.

First, while the statistical models created as part of this study are statistically significant, the explanatory power of our variables for different types of financial performance measures is relatively low (see adjusted R^2 in Table 3.7). As a result, the study's findings — that imitation leads to improved performance and that early-mover behaviour results in adverse performance outcomes — might only inform managers and executives to a small degree when making strategic decisions in the contexts of innovation, uncertainty and sustainability.

Second, the objective of the study was to determine whether imitation at an aggregated level leads to improved performance and consequently resolve friction within management research. While our contribution helps to advance management research, the general indication of a positive imitation-performance relationship can not be used

generally to make individual strategic decisions in specific situations to overcome concrete challenges.

Third, despite the influence of statistical and theoretical limitations on concrete strategic decisions, our study reveals a major finding relevant for practice:

Managers should not reject imitation as a strategy per se. Keeping imitation strategies among one's options might be the enabler of improved firm performance.

Furthermore, in light of our post-hoc findings - that firms being an early-mover in addressing sustainability topics as well as addressing sustainability in negatively connotated contexts experience positive performance outcomes - we advocate that:

Managers should engage and address sustainability contexts to realise improved financial performance and contribute to a more sustainable world.

5.3 Limitations

In the following, we extend the quality discussion and methodological choices made in the data and variable section of the empirical study. We highlight the limitations of our empirical study, focusing on (1) the empirical data concerning the annual reports and (2) the methodological choices centring around LDA topic modelling, and our variables of imitation, early-mover behaviour and firm performance.

First, our empirical data is based on annual reports, which introduces limitations. These include the timing of publication, the nature of the documents themselves, and the fact that companies use multiple channels to communicate their strategic decisions. Companies usually publish the annual report after the end of the company's fiscal year. In most cases, this fiscal year broadly follows a normal calendar year; however, certain companies have differing fiscal years, which could introduce potential time-lag effects. Another limitation noted in relation to 10-K forms is that the sentiment can both be forward-looking or backwards-looking (Bayer et al., 2017). This is a major limitation of our study, as we do not distinguish if an expressed sentence describes events from the past or the future. Lastly, the 10-K report is not the only source of communication companies use to disclose information (DeFond et al., 2005).

Second, our methodological choices of LDA topic modelling, machine-learning sentiment analysis, and variable design are the source of further limitations.

LDA topic modelling comes with the limitations of potential overlaps between topics, weakening the dictionary-based approach. Dictionaries themselves may contain

inherently positive or negative words, which could potentially introduce bias in the sentiment analysis. Another source of bias in LDA topic modelling is human interpretation of the output topics. For instance, identifying dictionaries like "innovation" or "sustainability" for each year relies on a subjective decision. This could potentially introduce bias and inter-observer risk from the person choosing the topics (Tirunillai and Tellis, 2014).

A limitation of sentiment analysis is the inherent statistical accuracy of the model. The model used, VADER, is furthermore not specifically trained to evaluate 10-K reports.

The variables developed as a measurement of imitation and early-mover behaviour are further limitations of our study as many factors could influence the accuracy of the variables. First, the variables are solely based on corporate communication (Form 10-K), which is only a proxy for the company's actual state and not executed actions or decisions taken. Second, the imitation and early-mover variables are supposed to capture their respective concepts; however, there might be instances not accounted for, where companies could score highly on a variable without actually being in the intended situation (e.g., imitator, early-mover, etc.). Moreover, the variable's focus on communication over specified time windows might not align perfectly with the actual timing of market impacts. Strategic moves could yield results outside of the fixed windows, and the benefits or disadvantages of being an early-mover or imitator might manifest long after initial communications and our longest 5-year outcome period. Lastly, these variables are limited by the underlying communication indicators and their own limitations. Despite carefully designing our measures and variables, the limitations caused by the very nature of our variable definition, and thus potential measurement errors, are hardly mitigable.

Lastly, a limitation of our financial performance measures is that, for example, hundreds of factors can influence stock performance (Harvey et al., 2016).

In summary, the study's limitations centre around the dependency on annual reports, which are subject to potential issues discussed above. Additionally, the methodological choices, including LDA topic modelling and sentiment analysis, introduce potential biases and inaccuracies in accurately capturing the nuances of corporate communication. Finally, the variables used to measure imitation and early-mover behaviour are constrained by their reliance on communication proxies and inherent design flaws.

5.4 Future Research

We argue for potential future research opportunities and thus contributions to a better understanding of the imitation-performance relationship by (1) expanding the dependent variables, (2) expanding the independent variables, and (3) utilising the positive indication on the nature of the imitation-performance relationship for qualitative research.

First, building on our methodological approach, we suggest an extension of the outcome variables to understand the broader impact of imitation behaviour on stakeholders and society. Future research could incorporate ESG and CSR variables measures alongside traditional financial metrics to provide a comprehensive understanding of imitation's implications on further stakeholders, the environment and society in general (Posen et al., 2023). On the other hand, sophisticating our findings via the implementation of imitation and early-mover variables into financial factor models (Eberhart et al., 2004; Fama and French, 1992, 1993) would allow for a better assessment of the financial impact and actual practical outcome of imitation behaviour. Such research would extend our findings beyond an indication of the nature of the imitation-performance relationship.

Second, we suggest extending the empirical data and independent variables to assess the imitation-performance relationship in different contexts and further increase the accuracy and generalisability of our findings. For example, altering the sample from S&P 500 companies to small-capitalisation companies would explore the role of imitation in emerging industries. On the other hand, keeping the S&P 500 but dissecting our study into industry clusters would potentially provide further insights into the nature of the imitation-performance relationship that currently goes unnoticed. Expanding the observation of firm sentiment from 10-K onto other means of communication, improving the measurement accuracy of our imitation and early-mover variables, and better understanding the theoretical implication of our sentiment breadth variable provide multiple avenues to build on our results.

Third, beyond the empirical and methodological opportunities, our robust assessment of the positive imitation-performance relationship offers guidance to return the imitation discussion to its qualitative origin. First and foremost, we suggest returning to Posen et al.'s (2023) discussion on the static and dynamic model of imitation and re-evaluating the underlying assumptions or exploring intra-organisational behaviour amidst our findings on the imitation-performance relationship. Such future research would help better implement and manage imitation strategies by shedding further light on why and how imitating firms derive their strategic imitation decisions.

Finally, we personally hope our post-hoc findings motivate the reader to further explore how imitation and early-mover behaviour in sustainability contexts can enhance firm performance and contribute to a greener future.

Bibliography

Argyres, N., L. Bigelow, and J. A. Nickerson (2015). "Dominant designs, innovation shocks, and the follower's dilemma." In: *Strategic Management Journal* 36 (2), pp. 216–234.

Banerjee, A. V. (1992). "A Simple Model of Herd Behavior." In: *The Quarterly Journal of Economics* 107 (3), pp. 797–817.

Banker, R., Z. Cao, N. M. Menon, and R. Mudambi (2013). "The Red Queen in action: The longitudinal effects of capital investments in the mobile telecommunications sector." In: *Industrial and Corporate Change* 22 (5), pp. 1195–1228.

Barney, J. (1991). "Firm Resources and Sustained Competitive Advantage." In: *Journal of Management* 17 (1), pp. 99–120.

Barreto, I. and C. Baden-Fuller (2006). "To Conform or To Perform? Mimetic Behaviour, Legitimacy-Based Groups and Performance Consequences*." In: *Journal of Management Studies* 43 (7), pp. 1559–1581.

Bayer, E., K. R. Tuli, and B. Skiera (2017). "Do Disclosures of Customer Metrics Lower Investors' and Analysts' Uncertainty but Hurt Firm Performance?" In: *Journal of Marketing Research* 54 (2), pp. 239–259.

Bell, E., A. Bryman, and B. Harley (2022). *Business research methods*. Oxford university press.

Bellstam, G., S. Bhagat, and J. A. Cookson (2021). "A Text-Based Analysis of Corporate Innovation." In: *Management Science* 67 (7), pp. 4004–4031.

BenMabrouk, H. (2018). "Cross-herding behavior between the stock market and the crude oil market during financial distress." In: *Managerial Finance* 44 (4), pp. 439–458.

Berk, J. and P. DeMarzo (2007). *Corporate Finance*. 6th ed. Pearson Education.

Berry, W. D. (1993). *Understanding regression assumptions*. Vol. 92. Sage.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." In: *Journal of Political Economy* 100 (5), pp. 992–1026.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003). "Latent Dirichlet Allocation." In: *Journal of Machine Learning Research* 3 (4/5), pp. 993–1022.

Bloom, N., T. A. Hassan, A. Kalyani, J. Lerner, and A. Tahoun (2021). *The Diffusion of New Technologies*.

Chen, C. D. and R. Demirer (2018). "The profitability of herding: evidence from Taiwan." In: *Managerial Finance* 44 (7), pp. 919–934.

Christensen, C. M. (1997). *The innovator's dilemma : when new technologies cause great firms to fail*. Harvard Business School Press, p. 252.

Christensen, C. M. and J. L. Bower (1996). "Customer Power, Strategic Investment, and the Failure of Leading Firms." In: *Strategic Management Journal* 17 (3), pp. 197–218.

Cohen, W. M. and D. A. Levinthal (1990). "Absorptive Capacity: A New Perspective on Learning and Innovation." In: *Administrative Science Quarterly* 35 (1), p. 128.

Creswell, J. W. and J. D. Creswell (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Daft, R. L. (1983). "Learning the craft of organizational research." In: *Academy of management review* 8.4, pp. 539–546.

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." In: *MIS Quarterly* 13 (3), p. 319.

Davis, J. P. and V. A. Aggarwal (2020). "Knowledge mobilization in the face of imitation: Microfoundations of knowledge aggregation and firm-level innovation." In: *Strategic Management Journal* 41 (11), pp. 1983–2014.

Deephouse, D. L. (1999). "To Be Different or to Be the Same? It's a Question (and a Theory) of Strategic Balance." In: *Strategic Management Journal* 20 (2).

DeFond, M. L., R. N. Hann, and X. Hu (2005). "Does the market value financial expertise on audit committees of boards of directors?" In: *Journal of accounting research* 43.2, pp. 153–193.

Delios, A., A. S. Gaur, and S. Makino (2008). "The Timing of International Expansion: Information, Rivalry and Imitation Among Japanese Firms, 1980–2002." In: *Journal of Management Studies* 45 (1), pp. 169–195.

Demirer, R. and H. Zhang (2019). "Do firm characteristics matter in explaining the herding effect on returns?" In: *Review of Financial Economics* 37 (2), pp. 256–271.

Devaraj, S., R. F. Easley, and J. M. Crant (2008). "Research Note —How Does Personality Matter? Relating the Five-Factor Model to Technology Acceptance and Use." In: *Information Systems Research* 19 (1), pp. 93–105.

Devenow, A. and I. Welch (1996). "Rational herding in financial economics." In: *European Economic Review* 40 (3-5), pp. 603–615.

DiMaggio, P. J., W. W. Powell, et al. (1983). "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields." In: *American sociological review* 48.2, pp. 147–160.

Dye, K. C., J. P. Eggers, and Z. Shapira (2014). "Trade-offs in a Tempest: Stakeholder Influence on Hurricane Evacuation Decisions." In: *Organization Science* 25 (4), pp. 1009–1025.

Eberhart, A. C., W. F. Maxwell, and A. R. Siddique (2004). "An examination of long-term abnormal stock returns and operating performance following R&D increases." In: *The journal of finance* 59.2, pp. 623–650.

Edmondson, A. C. and S. E. Mcmanus (2007). "Methodological fit in management field research." In: *Academy of Management Review* 32 (4), pp. 1246–1264.

Eisenhardt, K. M. and J. A. Martin (2000). "Dynamic capabilities: what are they?" In: *Strategic Management Journal* 21 (10-11), pp. 1105–1121.

Eisenhardt, K. M. and D. N. Sull (2001). "Strategy as simple rules." In: *Harvard Bus Pub* 6 (4), pp. 107–116.

Escrig-Olmedo, E., M. Fernández-Izquierdo, I. Ferrero-Ferrero, J. Rivera-Lirio, and M. Muñoz-Torres (2019). "Rating the Raters: Evaluating how ESG Rating Agencies Integrate Sustainability Principles." In: *Sustainability* 11 (3), p. 915.

Fama, E. F. and K. R. French (1992). "The Cross-Section of Expected Stock Returns." In: *The Journal of Finance* 47 (2), pp. 427–465.

Fama, E. F. and K. R. French (1993). "Common risk factors in the returns on stocks and bonds." In: *Journal of Financial Economics* 33 (1), pp. 3–56.

Fleming, L., R. P. Rumelt, D. E. Schendel, and D. J. Teece (1996). "Fundamental Issues in Strategy: A Research Agenda." In: *Administrative Science Quarterly* 41 (1), p. 196.

Fung, W. and D. A. Hsieh (1997). "Survivorship Bias and Investment Style in the Returns of CTAs." In: *The Journal of Portfolio Management* 24 (1), pp. 30–41.

Giachetti, C., J. Lampel, and S. L. Pira (2017). "Red Queen Competitive Imitation in the U.K. Mobile Phone Industry." In: *Academy of Management Journal* 60 (5), pp. 1882–1914.

Giachetti, C. and S. L. Pira (2022). "Catching up with the market leader: Does it pay to rapidly imitate its innovations?" In: *Research Policy* 51 (5), p. 104505.

Gill, J. and P. Johnson (2010). *Research Methods for Managers*. Sage.

Guerrero, M. and D. S. Siegel (2024). "Schumpeter meets Teece: Proposed metrics for assessing entrepreneurial innovation and dynamic capabilities in entrepreneurial ecosystems in an emerging economy." In: *Research Policy* 53 (5), p. 104984.

Guillén, M. F. (2003). "Experience, imitation, and the sequence of foreign entry: wholly owned and joint-venture manufacturing by South Korean firms and business groups in China, 1987–1995." In: *Journal of International Business Studies* 34 (2), pp. 185–198.

Harvey, C. R., Y. Liu, and H. Zhu (2016). " and the Cross-Section of Expected Returns." In: *Review of Financial Studies* 29 (1), pp. 5–68.

Haunschild, P. R. and A. S. Miner (1997). "Modes of Interorganizational Imitation: The Effects of Outcome Salience and Uncertainty." In: *Administrative Science Quarterly* 42 (3), p. 472.

Haveman, H. A. (1993). "Follow the Leader: Mimetic Isomorphism and Entry Into New Markets." In: *Administrative Science Quarterly* 38 (4), p. 593.

Henisz, W. J. and A. Delios (2001). "Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996." In: *Administrative Science Quarterly* 46 (3), pp. 443–475.

Hoppe, H. C. (2000). "Second-mover advantages in the strategic adoption of new technology under uncertainty." In: *International Journal of Industrial Organization* 18 (2), pp. 315–338.

Hutto, C. and E. Gilbert (2014). "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." In: *Proceedings of the International AAAI Conference on Web and Social Media* 8 (1), pp. 216–225.

Hwang, S. and M. Salmon (2004). "Market stress and herding." In: *Journal of Empirical Finance* 11 (4), pp. 585–616.

Itaoka, K. (2012). *Regression and interpretation low R-squared! Social Research Network 3 nd Meeting Noosa*.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." In: *The Quarterly Journal of Economics* 108 (3), pp. 577–598.

Jaffe, A. B., M. Trajtenberg, and M. S. Fogarty (2000). "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors." In: *American Economic Review* 90 (2), pp. 215–218.

Jiang, H. and M. Verardo (2013). "Does Herding Behavior Reveal Skill? An Analysis of Mutual Fund Performance." In: *SSRN Electronic Journal*.

Keister, L. A. (2004). "Capital Structure in Transition: The Transformation of Financial Strategies in China's Emerging Economy." In: *Organization Science* 15 (2), pp. 145–158.

Keynes, J. M. (1936). *The General Theory of Employment, Interest, and Money*. Cambridge University Press.

Knott, A. M. and S. F. Turner (2019). "An Innovation Theory of Headquarters Value in Multibusiness Firms." In: *Organization Science* 30 (1), pp. 19–39.

Komalasari, P. T., M. Asri, B. M. Purwanto, and B. Setiyono (2022). "Herding behaviour in the capital market: What do we know and what is next?" In: *Management Review Quarterly* 72 (3), pp. 745–787.

Korn, H. J. and J. A. C. Baum (1999). "Chance, Imitative, and Strategic Antecedents to Multimarket Contact." In: *Academy of Management Journal* 42 (2), pp. 171–193.

Lee, G. K., J. Lampel, and Z. Shapira (2020). "After the Storm Has Passed: Translating Crisis Experience into Useful Knowledge." In: *Organization Science* 31 (4), pp. 1037–1051.

Levinthal, D. A. (1997). "Adaptation on Rugged Landscapes." In: *Management Science* 43 (7), pp. 934–950.

Levinthal, D. A. (2011). "A behavioral approach to strategy—what's the alternative?" In: *Strategic Management Journal* 32 (13), pp. 1517–1523.

Levitt, B. and J. G. March (1988). "Organizational Learning." In: *Annual Review of Sociology* 14 (1), pp. 319–338.

Li, F., R. Lundholm, and M. Minnis (2013). "A Measure of Competition Based on 10-K Filings." In: *Journal of Accounting Research* 51 (2), pp. 399–436.

Lieberman, M. B. and D. B. Montgomery (1988). "First-mover advantages." In: *Strategic Management Journal* 9 (S1), pp. 41–58.

Ligorio, L., A. Venturelli, and F. Caputo (2022). "Tracing the boundaries between sustainable cities and cities for sustainable development. An LDA analysis of management studies." In: *Technological Forecasting and Social Change* 176, p. 121447.

Lippman, S. A. and R. P. Rumelt (1982). "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency under Competition." In: *The Bell Journal of Economics* 13 (2), p. 418.

Malone, T. W., P. Weill, R. K. Lai, V. T. D'Urso, G. Herman, T. G. Apel, and S. Woerner (2006). "Do Some Business Models Perform Better than Others?" In: *SSRN Electronic Journal*.

McDonald, R. M. and K. M. Eisenhardt (2020). "Parallel Play: Startups, Nascent Markets, and Effective Business-model Design." In: *Administrative Science Quarterly* 65 (2), pp. 483–523.

Meuleman, B., G. Loosveldt, and V. Emonds (2013). "Regression analysis: Assumptions and diagnostics." In: ed. by C. Wolf and H. Best. SAGE Publications Ltd, pp. 83–110.

Miller, C. C. and R. A. McKee (2021). "Decision comprehensiveness and the outcomes of firms: Reinterpreting and extending a recent meta-analysis." In: *Strategic Organization* 19 (3), pp. 441–448.

Miric, M., N. Jia, and K. G. Huang (2023). "Using supervised machine learning for large-scale classification in management research: The case for identifying artificial intelligence patents." In: *Strategic Management Journal* 44 (2), pp. 491–519.

Al-Natour, S. and O. Turetken (2020). "A comparative assessment of sentiment analysis and star ratings for consumer reviews." In: *International Journal of Information Management* 54, p. 102132.

Nauhaus, S., J. Luger, and S. Raisch (2021). "Strategic decision making in the digital age: Expert sentiment and corporate capital allocation." In: *Journal of Management Studies* 58.7, pp. 1933–1961.

Nguyen, H. and A. Agrawal (2023). "Disruptive Innovation and IPO Outcomes: Evidence from Machine Learning." In: *SSRN Electronic Journal*.

Nofsinger, J. R. and R. W. Sias (1999). "Herding and Feedback Trading by Institutional and Individual Investors." In: *The Journal of Finance* 54 (6), pp. 2263–2295.

Nousiainen, E., M. Ranta, M. Ylinen, and M. Järvenpää (2022). "Using Machine Learning and 10-K Filings to Measure Innovation." In: *SSRN Electronic Journal*.

Peteraf, M. A. (1993). "The cornerstones of competitive advantage: A resource-based view." In: *Strategic Management Journal* 14 (3), pp. 179–191.

Posen, H. E. and D. A. Levinthal (2012). "Chasing a moving target: Exploitation and exploration in dynamic environments." In: *Management science* 58.3, pp. 587–601.

Posen, H. E., J. Lee, and S. Yi (2013). "The power of imperfect imitation." In: *Strategic Management Journal* 34 (2), pp. 149–164.

Posen, H. E. and D. Martignoni (2018). "Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity." In: *Strategic Management Journal* 39 (5), pp. 1350–1369.

Posen, H. E., J. M. Ross, B. Wu, S. Benigni, and Z. Cao (2023). "Reconceptualizing Imitation: Implications for Dynamic Capabilities, Innovation, and Competitive Advantage." In: *Academy of Management Annals* 17 (1), pp. 74–112.

Raafat, R. M., N. Chater, and C. Frith (2009). "Herding in humans." In: *Trends in Cognitive Sciences* 13 (10), pp. 420–428.

Ramli, I., S. Agoes, and I. R. Setyawan (2015). "Information Asymmetry And The Role Of Foreign Investors In Daily Transactions During The Crisis; A Study Of Herding In The Indonesian Stock Exchange." In: *Journal of Applied Business Research (JABR)* 32 (1), p. 269.

Ribeiro, F. N., M. Araújo, P. Gonçalves, M. A. Gonçalves, and F. Benevenuto (2016). "SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods." In: *EPJ Data Science* 5 (1), p. 23.

Rivkin, J. W. (2000). "Imitation of Complex Strategies." In: *Management Science* 46 (6), pp. 824–844.

Robinson, T. R., E. Henry, and M. A. Broihahn (2012). *International Financial Statement Analysis*. 4th ed. John Wiley Sons Inc.

Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem." In: *The American Economic Review* 63 (2), pp. 134–139.

Rumelt, R. P. (1984). "Towards a strategic theory of the firm." In: *Competitive Strategic Management* 26 (3), pp. 556–570.

SP 500® Equity (2024).

Samba, C., P. Tabesh, I. C. Thanos, and V. M. Papadakis (2021). "Method in the madness? A meta-analysis on the strategic implications of decision comprehensiveness." In: *Strategic Organization* 19 (3), pp. 414–440.

Saunders, M., P. Lewis, and A. Thornhill (2009). *Research methods for business students*. Pearson education.

Scharfstein, D. S. and J. C. Stein (1990). "Herd Behavior and Investment." In: *The American Economic Review* 80 (3), pp. 465–479.

Schepman, A. and P. Rodway (2023). "The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust." In: *International Journal of Human–Computer Interaction* 39 (13), pp. 2724–2741.

Sharapov, D. and J.-M. Ross (2023). "Whom should a leader imitate? Using rivalry-based imitation to manage strategic risk in changing environments." In: *Strategic Management Journal* 44 (1), pp. 311–342.

Simon, D. H. and M. B. Lieberman (2010). "Internal and external influences on adoption decisions in multi-unit firms: the moderating effect of experience." In: *Strategic Organization* 8 (2), pp. 132–154.

Singh, V. (2013). "Did institutions herd during the internet bubble?" In: *Review of Quantitative Finance and Accounting* 41 (3), pp. 513–534.

Strang, D. and S. A. Soule (1998). "Diffusion in Organizations and Social Movements: From Hybrid Corn to Poison Pills." In: *Annual Review of Sociology* 24 (1), pp. 265–290.

Straub, D. and D. Gefen (2004). "Validation Guidelines for IS Positivist Research." In: *Communications of the Association for Information Systems* 13.

Suarez, F. and G. Lanzolla (2005). "The Half-Truth of First-Mover Advantage." In: *Harvard Business Review*.

Teece, D. J., G. Pisano, and A. Shuen (1997). "Dynamic capabilities and strategic management." In: *Strategic Management Journal* 18 (7), pp. 509–533.

Terlaak, A. and Y. Gong (2008). "Vicarious Learning And Inferential Accuracy in Adoption Processes." In: *Academy of Management Review* 33 (4), pp. 846–868.

Tirunillai, S. and G. J. Tellis (2014). "Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation." In: *Journal of Marketing Research* 51 (4), pp. 463–479.

Torres-Reyna, O. (2007). *Panel Data Analysis Fixed and Random Effects using Stata*.

Uotila, J., M. Maula, T. Keil, and S. A. Zahra (2009). "Exploration, exploitation, and financial performance: analysis of Samp;P 500 corporations." In: *Strategic Management Journal* 30 (2), pp. 221–231.

Vasudeva, G. (2013). "Weaving Together the Normative and Regulative Roles of Government: How the Norwegian Sovereign Wealth Fund's Responsible Conduct Is Shaping Firms' Cross-Border Investments." In: *Organization Science* 24 (6), pp. 1662–1682.

Wang, L., B. Wu, C. Pechmann, and Y. Wang (2023). "The performance effects of creative imitation on original products: Evidence from lab and field experiments." In: *Strategic Management Journal* 44 (1), pp. 171–196.

Wilcox, R. R. (2012). *Introduction to robust estimation and hypothesis testing*. Academic press.

Wolf, C. and H. Best (2013). *The SAGE Handbook of Regression Analysis and Causal Inference*. SAGE Publications Ltd.

Young, S. L. and J. B. Barney (2018). "Strategic Imitation." In: Palgrave Macmillan UK, pp. 1612–1617.

Yu, C.-M. J. and K. Ito (1988). "Oligopolistic Reaction and Foreign Direct Investment: The Case of the U.S. Tire and Textiles Industries." In: *Journal of International Business Studies* 19 (3), pp. 449–460.

Zhou, K. Z., C. K. (Yim, and D. K. Tse (2005). "The Effects of Strategic Orientations on Technology- and Market-Based Breakthrough Innovations." In: *Journal of Marketing* 69 (2), pp. 42–60.

Zhou, R. T. and R. N. Lai (2009). "Herding and information based trading." In: *Journal of Empirical Finance* 16 (3), pp. 388–393.

This document was typeset by the authors using LaTeX and builds on the Creative Commons CC BY 4.0 licensed template for PhD dissertations following the regulations of Stockholm School of Economics created by Adam Altmejd and uploaded to Overleaf by Nurit Nobel (https://www.overleaf.com/latex/templates/template-for-doctoral-dissertations-at-stockholm-school-of-economics/zwnndmctgwfd).

Appendix

OLS Regression Types

Pooled OLS regressions assume that there are no unobserved entity-specific or time-specific effects that correlate with the dependent variables. The choice to use pooled OLS depends on whether the data appears to be homogenous across entities and time, thus company A is expected to behave like company B, and year Y is similar to year Y+1. On the other hand, a fixed effects model is used when the unobserved heterogeneity is correlated with independent variables. By controlling for these specific effects, fixed-effects models provide more accurate estimates of the relationships between the independent and dependent variables. The decision for a fixed effects model furthermore includes the decision between fixed entity effects, fixed time effects, or both. Fixed entity effects are included when there is variation between entities, for example, different companies, that might affect the dependent variable. This model controls for all entity-specific, time-invariant characteristics. On the other hand, fixed time effects models control for shocks that affect all entities in the same way over a period of time, such as economic crises, wars, or policy changes. A model with both fixed entity and fixed time effects is used when both entity-specific characteristics and time-specific shocks are believed to influence the dependent variable. Lastly, a random effects model is suitable if the unobserved effects are assumed to be uncorrelated with the independent variables across all time periods and entities. It allows for generalisation of the results, yet is vulnerable to omitted variable bias.

LDA Dictionaries

Innovation

Year	Words
2009	technology, software, solution, network, client, segment, application, offer, consumer, support, high, design, offering, communication, platform, marketing, electronic, strategy, channel, device, provider, growth, digital, processing, computer, quality, global, license, enterprise, wireless, access, internet, competitive, area, range, advertising, need, retail, medium, drive, user, designed, relationship, sell, brand, online, hardware, storage, direct, many, capability, integrated, professional, leading, world, variety, focus, electronics, opportunity, video, local, help, site, infrastructure, retailer, live, third_party, competitor, large, tool, key, mobile, providing, delivery, core, distributor, manufacturer, category, organization, portfolio, resource, team, worldwide, wide, office, competition, compete, component, across, broad, feature, demand, single, entertainment, vendor, environment, develop, major, base, source
2010	network, application, offer, high, solution, communication, offering, global, platform, design, digital, strategy, support, channel, growth, provider, enterprise, wireless, access, medium, game, entertainment, area, online, device, integrated, electronic, internet, user, consumer, range, need, capability, quality, designed, video, opportunity, drive, help, world, local, subscriber, focus, core, leading, mobile, people, programming, many, competitive, tool, satellite, key, marketing, large, television, cable, feature, demand, infrastructure, semiconductor, portfolio, storage, advertising, wide, across, site, relationship, improve, variety, worldwide, providing, single, strategic, electronics, base, team, advanced, broad, enable, model, environment, competition, direct, competitor, web, center, home, expand, enhance, develop, small, focused, resource, organization, pc, major, more_than, live, film
2011	software, solution, network, communication, application, digital, device, hardware, wireless, enterprise, electronic, online, medium, support, provider, internet, user, storage, access, license, mobile, computer, platform, advertising, video, game, semiconductor, entertainment, cable, television, subscription, offering, programming, com, channel, pc, offer, site, high, electronics, infrastructure, processing, design, integrated, consumer, satellite, circuit, film, ip, web, local, computing, spectrum, subscriber, advanced, live, feature, automation, website, professional, search, networking, delivery, speed, woman, radio, phone, functionality, voice, education, marketing, telephone, display, carrier, interactive, licensing, distributor, tool, station, traffic, vendor, end_user, upgrade, release, technical, center, home, interface, manufacturer, power, training, maintenance, operator, consulting, developer, installation, imaging, designed, capacity, area

Continued on next page

Table 1 – continued from previous page

Year	Words
2012	solution, software, brand, support, application, consumer, enterprise, offer, offering, design, high, storage, competitive, supply, hardware, strategy, server, portfolio, platform, infrastructure, quality, range, global, marketing, need, personal, area, pc, large, integrated, networking, device, channel, designed, provider, cloud, model, focus, help, across, capability, competitor, online, electronic, opportunity, sell, many, leading, single, environment, key, relationship, demand, broad, tool, drive, center, user, semiconductor, strategic, computer, improve, delivery, manage, variety, wide, direct, world, focused, inventory, develop, network, retail, experience, distributor, game, partner, compete, computing, major, chain, enable, competition, retailer, access, source, outsourcing, manufacturing, worldwide, small, location, manufacturer, geographic, multi, team, core, consulting, medium, supplier, innovation
2013	client, solution, software, support, offer, global, offering, strategy, application, enterprise, high, platform, opportunity, portfolio, need, designed, help, focus, key, across, professional, range, model, capability, provider, team, infrastructure, relationship, competitive, strategic, quality, area, integrated, design, marketing, core, large, center, storage, improve, function, operational, providing, cloud, tool, leading, environment, many, drive, manage, organization, enhance, partner, single, enable, develop, experience, broad, research, base, delivery, channel, wide, focused, access, processing, expand, efficiency, work, subscription, deliver, manager, small, variety, strong, project, proprietary, competitor, developed, training, effort, outsourcing, serve, multiple, expect, direct, multi, world, critical, structure, objective, allows, computing, advantage, consulting, address, office, typically, them, expertise

Continued on next page

Table 1 – continued from previous page

Year	Words
2014	solution, support, client, software, application, offer, growth, offering, high, strategy, platform, need, range, global, key, provider, opportunity, across, infrastructure, portfolio, design, focus, strategic, user, enterprise, capability, competitive, electronic, device, direct, designed, allocation, center, quality, hardware, help, channel, organization, large, function, area, providing, develop, enable, wide, processing, network, research, cloud, integrated, operational, model, relationship, base, people, many, environment, improve, tool, professional, marketing, focused, core, manage, experience, storage, communication, broad, nrc, access, team, leading, deliver, expect, mobile, multiple, target, directly, drive, variety, delivery, expand, maintain, enhance, small, structure, computer, strong, effort, project, overall, serve, single, individual, multi, training, trend, installation, consulting, digital
2015	solution, software, network, support, offering, application, offer, high, enterprise, platform, digital, design, communication, growth, provider, device, cloud, strategy, mobile, channel, marketing, hardware, competitive, integrated, opportunity, global, user, capability, online, consumer, need, infrastructure, range, medium, access, portfolio, across, internet, video, quality, designed, key, experience, storage, delivery, area, help, advertising, focus, demand, research, center, electronic, leading, professional, competitor, strategic, subscription, cable, large, team, relationship, providing, many, competition, advanced, model, compete, wireless, local, improve, semiconductor, base, television, develop, enable, tool, deliver, programming, computer, partner, broad, direct, environment, expand, multiple, entertainment, focused, processing, world, people, drive, variety, expect, wide, feature, website, small, technical, manage

Continued on next page

Table 1 – continued from previous page

Year	Words
2016	solution, client, support, software, application, offering, growth, global, enterprise, design, high, offer, strategy, platform, cloud, vehicle, range, across, need, capability, quality, opportunity, competitive, key, experience, focus, center, delivery, integrated, portfolio, hardware, strategic, infrastructure, help, designed, provider, marketing, leading, many, large, improve, demand, device, area, professional, team, supply, enable, model, relationship, environment, storage, broad, tool, channel, world, automotive, subscription, wide, drive, variety, processing, focused, deliver, chain, competitor, efficiency, providing, core, technical, partner, aerospace, manage, expect, base, develop, electronic, industrial, research, strong, innovation, expand, enhance, distributor, operational, function, single, organization, source, people, sell, multiple, engineering, highly, critical, trend, supplier, direct, location, access
2017	solution, client, application, offering, support, offer, network, platform, enterprise, provider, high, software, device, across, cloud, design, capability, range, area, strategy, portfolio, help, growth, designed, focus, need, leading, enable, mobile, communication, user, infrastructure, marketing, quality, access, competitive, broad, organization, contact, integrated, electronic, world, relationship, model, channel, providing, opportunity, people, key, experience, competitor, compete, large, direct, wide, center, team, focused, tool, strategic, digital, professional, resort, improve, partner, processing, deliver, multi, competition, variety, many, single, research, develop, work, serve, drive, innovation, manage, advanced, science, environment, internet, expand, core, data_center, expertise, address, around, storage, enhance, geographic, function, operational, small, highly, delivery, medium, integration, feature

Continued on next page

Table 1 – continued from previous page

Year	Words
2018	technology, solution, segment, software, network, client, digital, consumer, offering, support, application, global, offer, platform, communication, high, enterprise, growth, strategy, provider, channel, marketing, cloud, believe, design, across, competitive, portfolio, capability, device, strategic, key, hardware, need, range, experience, opportunity, mobile, focus, access, leading, online, quality, area, subscription, help, medium, research, infrastructure, center, professional, integrated, designed, partner, large, relationship, model, competitor, world, many, advanced, home, license, user, electronic, processing, competition, compete, improve, people, advertising, direct, category, broad, providing, brand, enable, resource, wide, geographic, variety, wireless, delivery, develop, team, innovation, focused, internet, tool, environment, multi, deliver, demand, video, local, drive, retailer, organization, small, multiple
2019	solution, support, offering, software, global, application, growth, high, strategy, platform, design, manufacturing, cloud, competitive, offer, across, vehicle, strategic, portfolio, supply, enterprise, device, capability, experience, range, quality, opportunity, key, demand, leading, provider, need, focus, model, research, semiconductor, help, center, network, team, integrated, hardware, infrastructure, many, world, designed, enable, large, manufacturer, delivery, area, work, tool, improve, competitor, innovation, marketing, user, advanced, environment, people, broad, professional, relationship, access, variety, competition, develop, distributor, deliver, compete, chain, providing, drive, wide, focused, expect, channel, single, base, effort, strong, multi, highly, communication, efficiency, critical, small, storage, multiple, production, source, processing, location, industrial, trend, serve, direct, core, electronic

Continued on next page

Table 1 – continued from previous page

Year	Words
2020	solution, support, digital, software, strategy, application, offering, cloud, high, across, design, offer, need, platform, key, strategic, opportunity, capability, portfolio, focus, enterprise, team, help, experience, range, infrastructure, designed, model, enable, integrated, leading, people, improve, work, quality, device, research, operational, semiconductor, environment, provider, competitive, focused, advanced, tool, drive, deliver, professional, organization, area, core, communication, providing, center, manage, efficiency, broad, network, strong, hardware, enhance, innovation, base, access, many, world, processing, large, initiative, delivery, effort, target, goal, wide, critical, specific, line, multi, serve, develop, storage, multiple, overall, scale, demand, individual, expertise, variety, decision, expect, ensure, best, structure, highly, expand, better, component, address, function, community
2021	store, solution, brand, digital, consumer, software, network, segment, manufacturing, retail, channel, offering, platform, application, cloud, marketing, supply, design, offer, device, high, license, food, production, distributor, advertising, semiconductor, vehicle, provider, inventory, center, manufacturer, merchandise, component, capability, online, across, sell, portfolio, quality, partner, mobile, rider, medium, location, direct, fiscal, communication, world, integrated, delivery, experience, chain, infrastructure, television, range, hardware, manufacture, category, leading, supplier, sold, commerce, programming, research, retailer, model, access, subscription, user, enterprise, line, competition, many, variety, large, sport, single, wireless, competitive, home, storage, deliver, competitor, small, good, licensing, relationship, wholesale, selling, innovation, area, country, tool, supply_chain, designed, electronic, processing, advanced, trademark
2022	solution, store, brand, software, digital, semiconductor, consumer, application, network, manufacturing, offering, platform, offer, channel, device, cloud, retail, high, design, marketing, supply, display, food, capability, portfolio, across, distributor, mobile, manufacturer, provider, quality, infrastructure, growth, online, range, chain, enable, license, delivery, direct, center, access, experience, research, world, segment, advertising, component, partner, medium, location, sell, advanced, leading, integrated, competitor, many, storage, large, supplier, subscription, communication, compete, production, tool, user, enterprise, competitive, need, hardware, inventory, electronic, relationship, broad, supply_chain, variety, category, deliver, designed, multiple, commerce, focus, help, single, area, strategy, competition, innovation, model, retailer, wireless, international, typically, key, improve, licensing, automation, restaurant, packaging, line

Sustainability

Year	Words
2009	environmental, gulf, site, recovery, storm, regulatory, hurricane, decommissioning, filing, emission, waste, water, proceeding, purchased_power, remediation, epa, approval, damage, decision, approved, rider, air, recover, permit, clean, proposed, formula, hazardous, expenditure, request, disposal, review, louisiana, plant, area, legislation, restoration, katrina, proposal, hurricane_katrina, july, substance, local, implemented, agency, implementation, reserve, application, final, interim, owner, project, responsible, regulated, must, staff, study, cannot, implement, contamination, protection, governmental, spent, jurisdiction, test, work, requires, phase, investigation, reduce, base, caused, response, treatment, nuclear_decommissioning, limit, comply, several, hearing, cleanup, timing, removal, usage, quality, recovered_through, mechanism, meet, initiative, source, predict, pollution, construction, discharge, address, co, assessment, belief, collection, petition, safety
2010	environmental, site, water, emission, remediation, safety, waste, compliance, hazardous, chemical, epa, air, work, substance, protection, asbestos, disposal, fire, expenditure, clean, greenhouse, environmental_matter, greenhouse_gas, permit, agency, health, contamination, government, climate, local, damage, owner, responsible, injury, material_adverse, environment, area, climate_change, cleanup, legislation, investigation, hazardous_substance, discharge, comply, plant, environmental_protection, closure, fine, exposure, cannot, environmental_remediation, personal_injury, potentially, quality, several, storage, governmental, source, pollution, possible, remedial, capital_expenditure, electrical, operator, release, handling, labor, response, regulated, industrial, bargaining, transportation, penalty, carbon, monitoring, incur, resource, treatment, stringent, construction, many, personal, collective, regulatory, assessment, belief, known, removal, contingency, collective_bargaining, identified, although, former, liable, trade, assurance, proceeding, range, recovery, proposed

Continued on next page

Table 2 – continued from previous page

Year	Words
2011	environmental, energy, power, project, site, plant, electric, water, utility, construction, emission, remediation, mining, generation, hazardous, fuel, chemical, coal, waste, gas, transmission, regulatory, mine, safety, epa, recovery, air, expenditure, electricity, substance, metal, capacity, regulated, ferc, disposal, industrial, clean, supply, storage, permit, mill, environmental_matter, contamination, greenhouse, nuclear, greenhouse_gas, ton, hazardous_substance, climate, transportation, approved, maintenance, source, carbon, resource, surface, responsible, wind, capital_expenditure, work, discharge, infrastructure, mineral, renewable, generating, environment, health, pound, environmental_remediation, environmental_protection, pollution, retail, owner, natural, tariff, electrical, climate_change, purchased, removal, mechanism, demand, treatment, production, underground, closure, fire, phase, probable, conservation, protection, area, recover, monitoring, several, land, weather, engineering, base, station, final
2012	environmental, site, water, generation, transmission, emission, remediation, safety, epa, waste, hazardous, disposal, ohio, substance, air, permit, legislation, clean, local, expenditure, agency, climate, proposed, wind, authority, environmental_matter, health, chemical, responsible, comply, climate_change, environment, protection, cleanup, discharge, storage, contamination, area, jurisdiction, work, wholesale, operator, fine, greenhouse, owner, pollution, hazardous_substance, greenhouse_gas, review, final, environmental_protection, response, several, national, closure, treatment, recovery, regulated, governmental, personal, injury, weather, investigation, plant, damage, material_adverse, source, impose, carbon, cannot, penalty, adopted, violation, potentially, recover, removal, construction, cook, stringent, limit, operate, personal_injury, capital_expenditure, monitoring, predict, liable, department, exposure, assessment, degree, incident, installation, incur, hazardous_material, initiative, quality, although, filing, regional, many

Continued on next page

Table 2 – continued from previous page

Year	Words
2013	environmental, site, water, air, nuclear, emission, epa, remediation, hazardous, waste, expenditure, substance, clean, damage, disposal, climate, cannot, greenhouse, greenhouse_gas, chemical, pollution, final, source, contamination, health, cleanup, climate_change, outcome, storm, permit, capital_expenditure, ultimate, proposed, quality, safety, discharge, area, environmental_protection, hazardous_substance, environmental_remediation, storage, remedial, responsible, environment, agency, owner, unreasonably, carbon, clean_air, removal, retirement, recovery, ultimate_outcome, renewable, study, injury, release, protection, hurricane, work, incident, monitoring, plant, response, installation, several, comply, substantial, regulated, statute, incur, global, potentially, caused, operated, numerous, belief, personal, replacement, personal_injury, solid, liable, project, published, depend, therefore, florida, identified, so, spent, operational, construction, handling, many, adopted, will_depend, revised, permitting, governmental, legislation
2014	environmental, site, remediation, epa, emission, water, air, coal, waste, contamination, hazardous, disposal, mining, plant, substance, clean, expenditure, safety, permit, source, area, proposed, agency, mine, quality, climate, protection, cannot, surface, health, discharge, comply, co, greenhouse, environmental_protection, final, pollution, local, greenhouse_gas, regulated, combustion, hazardous_substance, storage, cleanup, ton, environment, environmental_remediation, responsible, capital_expenditure, climate_change, monitoring, chemical, release, generating, clean_air, river, response, investigation, residual, so, installation, several, limit, asbestos, operator, owner, mineral, ultimate, outcome, legislation, incident, solid, closure, refueling, clean_up, pollutant, handling, damage, cooling, fuel, assessment, national, incur, potentially, substantial, accident, stringent, operated, work, injury, range, structure, treatment, air_quality, past, governmental, study, violation, removal, located

Continued on next page

Table 2 – continued from previous page

Year	Words
2015	environmental, site, epa, emission, water, air, proposed, remediation, waste, clean, hazardous, disposal, safety, final, expenditure, co, permit, substance, cannot, area, source, plant, quality, limit, comply, climate, agency, discharge, pollution, contractor, closure, capital_expenditure, responsible, protection, contamination, owner, environmental_remediation, monitoring, asbestos, study, storage, hazardous_substance, substantial, damage, health, cleanup, clean_air, release, assessment, ultimate, construction, environmental_protection, operator, environment, several, implementation, so, response, legislation, local, installation, published, regulated, revised, incident, requires, statute, guideline, chemical, national, outcome, inspection, generating, treatment, ongoing, solid, adopted, injury, clean_up, air_quality, work, enforcement, stringent, remedial, proposal, hour, penalty, operational, revision, investigation, potentially, impose, known, depend, modified, significantly, range, violation, consent, structure
2016	environmental, nuclear, site, emission, plant, air, water, epa, wholesale, remediation, hazardous, waste, final, clean, disposal, permit, expenditure, substance, area, safety, fuel, co, cannot, climate, spent, coal, storage, generating, agency, quality, source, energy_efficiency, discharge, closure, capital_expenditure, gulf, license, proposed, owner, hazardous_substance, decision, contamination, comply, greenhouse, operator, response, implementation, final_rule, study, legislation, pollution, ultimate, greenhouse_gas, removal, environmental_remediation, carbon, responsible, limit, damage, construction, review, cleanup, revised, climate_change, environment, challenge, outcome, missouri, commodity, protection, monitoring, regulated, timing, health, mechanism, so, environmental_protection, national, cap, address, published, fired, pollutant, installation, local, clean_up, several, conservation, court, inspection, point, design, treatment, finalized, annual_report, gas, release, statute, landfill, clean_air

Continued on next page

Table 2 – continued from previous page

Year	Words
2017	environmental, site, coal, emission, mining, water, safety, epa, remediation, mine, waste, air, hazardous, plant, disposal, chemical, permit, substance, clean, agency, final, proposed, health, contamination, expenditure, local, comply, climate, supreme, review, legislation, supreme_court, surface, discharge, work, protection, area, government, limit, quality, source, climate_change, closure, several, ash, penalty, environment, operator, river, hazardous_substance, response, responsible, requires, implementation, owner, greenhouse, pollution, co, environmental_protection, department, assessment, greenhouse_gas, carbon, investigation, cannot, impose, asbestos, regulated, stringent, ton, potentially, adopted, storage, national, study, must, monitoring, cleanup, range, construction, land, permitting, injury, environmental_remediation, basin, fine, governmental, capital_expenditure, treatment, release, operate, so, violation, approval, address, phase, regulate, cap, handling, predict
2018	environmental, water, ferc, site, emission, air, epa, remediation, waste, hazardous, clean, disposal, el, area, permit, asbestos, coal, safety, pipe, steam, chemical, substance, fossil, agency, discharge, climate, bore, iso, steel, module, electrical, ghg, hazardous_substance, source, conduit, package, contamination, carbon, greenhouse, expenditure, protection, co, responsible, local, greenhouse_gas, quality, fired, generator, battery, closure, final, climate_change, comply, nation, work, pollution, generating, storage, environmental_remediation, cleanup, monitoring, stipulation, environment, annex, pump, resource, surface, national, rulemaking, room, clean_up, response, ton, solar, wall, removal, landfill, dioxide, vessel, pollutant, stay, treatment, coal_fired, legislation, permitting, potentially, circuit, specific, requires, regional, several, combustion, handling, capital_expenditure, approving, stringent, release, proposed, finalized, published

Continued on next page

Table 2 – continued from previous page

Year	Words
2019	environmental, site, water, emission, epa, remediation, coal, safety, air, waste, hazardous, permit, clean, ash, plant, disposal, mine, climate, substance, mining, closure, proposed, local, area, climate_change, agency, final, expenditure, operator, chemical, ccr, contamination, source, discharge, carbon, comply, co, quality, surface, legislation, greenhouse, hazardous_substance, health, river, response, greenhouse_gas, protection, responsible, environment, national, pollution, several, monitoring, hearing, requires, approval, fired, generating, project, sustainability, environmental_remediation, remedial, permitting, storage, capital_expenditure, government, so, owner, study, review, implementation, cleanup, regional, release, work, cannot, adopted, arkansas, impose, regulated, stringent, operate, treatment, effort, located, address, steam, challenge, department, limit, finalized, installation, potentially, handling, improvement, predict, specific, rulemaking, design, transport
2020	environmental, recovery, emission, water, ferc, approved, site, waste, epa, air, remediation, base, proposed, filing, virginia, formula, refund, approval, plant, september, clean, recovered, hazardous, review, final, october, disposal, april, mechanism, november, proceeding, regulated, recover, permit, expenditure, illinois, climate, july, legislation, application, request, substance, ghg, august, carbon, agency, project, tariff, greenhouse, ash, climate_change, response, discharge, allowed, staff, decision, greenhouse_gas, structure, closure, implementation, annually, collected, outcome, jersey, coal, local, efficiency, cannot, ratemaking, new_jersey, hazardous_substance, pending, requires, removal, billing, capital_expenditure, surcharge, refunded, cap, study, revised, several, pollution, source, requested, test, contamination, area, co, incentive, south_carolina, environmental_remediation, implemented, bill, expects, kentucky, timing, energy_efficiency, quality, implement

Continued on next page

Table 2 – continued from previous page

Year	Words
2021	environmental, site, emission, water, climate, safety, waste, disposal, remediation, air, hazardous, climate_change, permit, carbon, medical_device, substance, local, mechanism, plant, expenditure, cell, inspection, quality, agency, area, ash, coal, source, greenhouse, chemical, ghg, comply, health, greenhouse_gas, project, discharge, closure, storage, contamination, environment, operator, responsible, protection, mine, legislation, requires, response, co, capital_expenditure, pollution, surface, landfill, several, grid, operate, construction, wholesale, design, limit, monitoring, produced, stringent, goal, study, governmental, sustainability, permitting, potentially, assessment, substantial, regional, mineral, national, impose, fuel, owner, cannot, land, address, review, environmental_remediation, low, modification, column, zero, solid, transportation, initiative, fossil, handling, many, installation, injury, administration, treatment, significantly, release, regulate, renewable, predict
2022	environmental, emission, water, site, air, climate, waste, carbon, remediation, safety, epa, climate_change, hazardous, disposal, clean, chemical, substance, permit, fuel, local, ghg, flight, greenhouse, ship, fossil, expenditure, ash, agency, area, greenhouse_gas, contamination, storage, quality, supreme, source, goal, environment, discharge, legislation, closure, health, co, ghg_emission, comply, zero, transportation, fleet, supreme_court, petition, ton, protection, gas, hazardous_substance, several, sustainability, response, monitoring, scope, project, target, proposed, capital_expenditure, pollution, review, regional, approval, national, operate, operator, sea, responsible, passenger, hazardous_material, requires, handling, cannot, international, limit, consumption, surface, address, governmental, operational, achieve, release, conservation, efficiency, administration, california, pollutant, environmental_remediation, stringent, potentially, inspection, solid, department, treatment, carrier, fine, owner

Uncertainty

Year	Words
2009	adversely, adverse, demand, failure, able, material_adverse, materially, many, competitor, competitive, supplier, regulatory, substantial, could_adversely, reduce, obtain, unable, government, maintain, delay, cannot, competition, intellectual_property, intellectual, third_party, significantly, assurance, negatively, country, adversely_affected, adversely_affect, decline, economic_condition, disruption, meet, growth, key, uncertainty, liquidity, fluctuation, availability, need, operate, difficult, global, resource, profitability, experience, although, limit, incur, personnel, difficulty, successfully, expect, consumer, harm, technology, compete, timely, fail, face, success, access, litigation, depends, retain, depend, source, anticipated, volatility, negative, restriction, dependent, particularly, recent, political, sufficient, there_can, adequate, develop, comply, successful, environment, local, economy, sell, impacted, effort, downturn, strategy, predict, supply, particular, adverse_effect, even, inability, example, manage, quality
2010	adversely, demand, competitive, failure, able, competitor, many, material_adverse, materially, could_adversely, substantial, unable, competition, growth, delay, government, negatively, regulatory, maintain, adversely_affect, obtain, significantly, adversely_affected, cannot, disruption, decline, supplier, meet, need, economic_condition, profitability, assurance, key, experience, resource, availability, fluctuation, difficult, uncertainty, operate, although, country, successfully, global, difficulty, personnel, compete, liquidity, limit, manufacturing, harm, success, incur, fail, face, particularly, timely, depends, depend, access, pricing, source, litigation, retain, relationship, negative, manage, expect, anticipated, dependent, develop, strategy, sufficient, highly, higher, environment, impacted, large, political, successful, there_can, predict, effectively, recent, inability, particular, quality, reputation, damage, even, volatility, lead, example, greater, adverse_effect, restriction, economy, experienced, supply, adequate

Continued on next page

Table 3 – continued from previous page

Year	Words
2011	adversely, demand, failure, able, could_adversely, competitor, material_adverse, materially, competitive, unable, delay, adversely_affect, many, supplier, negatively, substantial, disruption, adversely_affected, maintain, obtain, competition, cannot, fluctuation, growth, decline, profitability, economic_condition, assurance, personnel, experience, resource, significantly, country, manufacturing, successfully, key, harm, difficulty, difficult, need, litigation, damage, availability, uncertainty, incur, operate, compete, success, timely, retain, fail, global, reputation, depend, develop, depends, face, although, negative, limit, anticipated, political, successful, inability, dependent, there_can, adverse_effect, sufficient, expect, particularly, impacted, effectively, source, access, government, relationship, manage, lead, highly, even, margin, higher, particular, supply, pricing, greater, liquidity, experienced, spending, effort, labor, disaster, adequate, large, rely, reduced, quality, volatility, materially_adversely, economy
2012	adversely, failure, demand, supplier, could_adversely, able, competitive, competitor, delay, disruption, many, material_adverse, country, unable, adversely_affected, adversely_affect, negatively, substantial, obtain, maintain, harm, competition, key, cannot, successfully, profitability, decline, litigation, experience, assurance, significantly, difficult, reputation, fail, margin, fluctuation, uncertainty, source, incur, compete, difficulty, damage, retain, operate, no_assurance, need, example, anticipated, economic_condition, pricing, effort, limit, personnel, manage, particularly, political, suffer, intellectual, intellectual_property, access, must, local, depend, availability, face, negative, although, dependent, timely, develop, adverse_effect, outside, effectively, depends, lead, rely, supply, particular, global, successful, protect, relationship, even, quality, there_can, restriction, comply, pressure, success, impacted, reduced, consumer, practice, jurisdiction, manufacturing, sufficient, sell, trade, inability, breach

Continued on next page

Table 3 – continued from previous page

Year	Words
2013	adversely, failure, demand, competitive, could_adversely, able, competitor, materially, material_adverse, unable, disruption, adversely_affect, negatively, delay, many, maintain, competition, adversely_affected, substantial, experience, decline, obtain, profitability, uncertainty, economic_condition, key, fluctuation, personnel, significantly, harm, successfully, reputation, difficult, incur, compete, assurance, difficulty, availability, cannot, operate, success, need, negative, retain, depends, supplier, anticipated, timely, face, damage, fail, develop, effectively, particularly, successful, access, impacted, relationship, dependent, manage, depend, global, political, pricing, consumer, effort, rely, inability, adverse_effect, although, sufficient, limit, volatility, disaster, highly, there_can, source, greater, strategy, spending, even, environment, sell, country, example, experienced, achieve, expect, economy, protect, particular, margin, materially_adversely, breach, reduced, expectation, pressure, lead, suffer, litigation
2014	adversely, adverse, adversely_affect, failure, material_adverse, demand, able, cause, materially, unable, delay, disruption, many, substantial, incur, government, competitive, cannot, operate, negatively, country, maintain, adversely_affected, competitor, uncertainty, obtain, damage, decline, experience, reputation, fail, significantly, growth, difficult, competition, assurance, harm, although, key, fluctuation, local, limit, timely, access, profitability, successfully, difficulty, retain, personnel, anticipated, comply, availability, negative, litigation, liquidity, need, political, lead, example, become, restriction, depends, higher, effectively, suffer, adverse_effect, inability, breach, impacted, expect, successful, supplier, sufficient, even, success, governmental, face, environment, particularly, dependent, global, rely, depend, interruption, outside, manage, particular, pricing, predict, past, manner, develop, effort, greater, labor, disaster, achieve, changing, strategic, increasing

Continued on next page

Table 3 – continued from previous page

Year	Words
2015	adversely, adverse, adversely_affect, failure, demand, material_adverse, able, materially, unable, delay, negatively, disruption, competitive, substantial, many, damage, maintain, adversely_affected, competitor, experience, incur, cannot, reputation, operate, decline, growth, supplier, obtain, significantly, key, assurance, access, limit, profitability, need, successfully, harm, difficulty, personnel, competition, difficult, availability, negative, fail, timely, fluctuation, uncertainty, retain, breach, litigation, higher, adverse_effect, depends, lead, manage, anticipated, source, although, depend, effectively, impacted, inability, face, example, successful, particularly, dependent, success, sufficient, labor, liquidity, disaster, political, even, environment, large, relationship, become, supply, attack, negatively_impact, rely, compete, develop, strategy, effort, pricing, particular, global, reduced, volatility, restriction, operational, challenge, increasing, materially_adversely, local, natural, achieve, past
2016	adversely, adversely_affect, failure, demand, material_adverse, materially, able, competitive, government, unable, disruption, supplier, delay, negatively, many, substantial, country, maintain, competitor, obtain, incur, adversely_affected, operate, uncertainty, reputation, cannot, competition, decline, growth, damage, experience, key, harm, access, fluctuation, assurance, successfully, significantly, availability, difficulty, negative, local, personnel, anticipated, fail, difficult, litigation, profitability, need, limit, timely, comply, example, face, retain, although, depends, supply, adverse_effect, impacted, source, breach, lead, successful, particularly, depend, restriction, success, effectively, effort, higher, inability, global, political, manage, develop, become, even, particular, dependent, compete, relationship, attack, labor, sell, negatively_impact, governmental, sufficient, practice, environment, rely, pricing, disaster, volatility, liquidity, quality, strategy, outside, expect, manner

Continued on next page

Table 3 – continued from previous page

Year	Words
2017	adversely, adversely_affect, failure, material_adverse, government, able, unable, disruption, delay, country, substantial, negatively, maintain, cannot, competitor, damage, reputation, obtain, comply, adversely_affected, harm, incur, many, operate, litigation, access, supplier, fail, assurance, competitive, breach, personnel, limit, experience, difficulty, restriction, jurisdiction, practice, successfully, retain, example, difficult, key, negative, timely, protection, lead, demand, governmental, although, face, local, uncertainty, decline, significantly, adverse_effect, relationship, attack, competition, inability, protect, there_can, rely, suffer, depend, penalty, trade, political, need, depends, successful, even, effectively, effort, negatively_impact, labor, anticipated, profitability, prevent, manner, privacy, success, sufficient, compete, conduct, outside, disaster, cyber, fluctuation, materially_adversely, develop, particular, measure, fine, interruption, manage, thing, sanction, impacted, them
2018	adversely, adverse, adversely_affect, failure, material_adverse, technology, government, demand, able, materially, disruption, unable, delay, reduce, maintain, country, substantial, many, negatively, competitive, reputation, competitor, damage, obtain, operate, access, limit, comply, cannot, incur, supplier, meet, adversely_affected, harm, significantly, experience, assurance, resource, successfully, key, decline, protection, competition, breach, restriction, example, practice, uncertainty, personnel, fail, negative, local, timely, growth, difficult, difficulty, lead, need, retain, jurisdiction, international, availability, attack, although, litigation, profitability, fluctuation, anticipated, protect, political, source, become, effort, face, adverse_effect, trade, consumer, relationship, governmental, penalty, successful, depends, impacted, particular, higher, inability, conduct, depend, sufficient, effectively, develop, rely, manage, measure, particularly, negatively_impact, cyber, even, operational, success

Continued on next page

Table 3 – continued from previous page

Year	Words
2019	adversely, adversely_affect, failure, material_adverse, supplier, demand, intellectual_property, intellectual, able, disruption, government, negatively, unable, delay, substantial, country, competitor, many, competitive, damage, reputation, obtain, operate, uncertainty, adversely_affected, trade, harm, access, incur, cannot, competition, assurance, personnel, breach, litigation, supply, international, decline, successfully, comply, key, experience, significantly, fluctuation, negative, fail, growth, difficulty, example, attack, timely, availability, anticipated, protect, difficult, retain, effort, political, profitability, lead, face, restriction, limit, labor, vendor, relationship, manufacturing, adverse_effect, protection, governmental, depends, local, impacted, negatively_impact, compete, source, successful, although, success, need, develop, raw_material, effectively, inability, become, raw, cyber, rely, depend, particularly, penalty, others, manage, disaster, global, sell, practice, dependent, sufficient, even
2020	adversely_affect, failure, pandemic, material_adverse, disruption, demand, negatively, government, able, delay, unable, supplier, many, substantial, reputation, country, restriction, operate, damage, uncertainty, cannot, access, incur, obtain, adversely_affected, comply, experience, negative, harm, decline, limit, assurance, example, significantly, impacted, competitive, key, personnel, successfully, measure, litigation, fail, attack, timely, difficult, availability, effort, practice, difficulty, governmental, retain, lead, local, need, breach, political, adverse_effect, negatively_impact, profitability, competitor, operational, inability, trade, labor, supply, manage, cyber, anticipated, effectively, face, although, depends, privacy, protection, fluctuation, depend, successful, relationship, there_can, penalty, additionally, particularly, source, higher, rely, develop, environment, disaster, challenge, liquidity, volatility, sufficient, health, manner, personal, even, predict, reduced, conduct, success

Continued on next page

Table 3 – continued from previous page

Year	Words
2021	adversely_affect, able, competitive, unable, negatively, supplier, material_adverse, failure, delay, competitor, key, experience, competition, successfully, obtain, many, adversely_affected, strategic, anticipated, manufacturing, strategy, substantial, availability, disruption, supply, timely, personnel, decline, operate, profitability, retain, incur, cannot, relationship, assurance, difficulty, depends, limit, fail, inability, effectively, negatively_impact, develop, uncertainty, significantly, face, difficult, compete, sufficient, success, lead, depend, source, achieve, sell, manage, pricing, negative, example, fluctuation, alternative, successful, dependent, higher, attract, access, favorable, lower, adverse_effect, highly, component, there_can, partner, challenge, quality, opportunity, manner, initiative, volatility, impacted, particular, particularly, liquidity, acquired, even, margin, increasing, expectation, seek, effort, although, changing, additionally, pressure, harm, labor, greater, large, enter, predict
2022	covid, pandemic, adversely_affect, supply, disruption, supplier, negatively, covid_pandemic, delay, manufacturing, unable, uncertainty, material_adverse, able, impacted, supply_chain, chain, competitive, availability, many, experience, labor, operate, growth, key, competitor, successfully, timely, negative, adversely_affected, substantial, decline, need, strategy, country, lead, fluctuation, effectively, cannot, political, incur, strategic, significantly, anticipated, increasing, profitability, retain, personnel, difficulty, measure, competition, difficult, obtain, experienced, health, challenge, volatility, manage, negatively_impact, reduced, liquidity, inability, depends, trade, consumer, production, fail, depend, relationship, source, operational, shortage, international, limit, response, highly, inventory, particularly, predict, although, higher, economy, assurance, component, governmental, sufficient, effort, ongoing, changing, environment, develop, access, reputation, face, sell, success, partner, timing, achieve, pressure

Note on Generative AI

We use generative AI for proof-reading. No other text generated by any generative AI tool has been copied, except for this very section in the appendix. The decision to abstain from using generative AI is rooted in several key concerns regarding the reliability and accuracy of AI-generated content:

1. Inability to verify outputs: Given the nature of generative AI, it is challenging to verify the correctness and appropriateness of every output it produces. This raises concerns

about the introduction of errors or unintended meanings, which could compromise the integrity of the thesis.

2. Lack of control over text modification: Even with precise prompts intended to refine or clarify text, generative AI tools often alter words or phrases in ways that were not requested. This lack of predictability in the tool's output undermines our ability to maintain the original intent and precision of the thesis content.

3. Ethical considerations: Ensuring that all content in this thesis is the result of original thought and rigorous academic research is paramount. Relying on AI-generated content could blur the lines between original work and machine-assisted output, which could raise ethical questions about the authorship and originality of the thesis.

To maintain the highest standards of academic integrity, the only instance where generative AI has been employed is in the drafting of this specific note. The prompt used to generate this note is included below, and the generated output has been carefully reviewed to ensure it aligns with the intended message and meets the thesis requirements.

The following prompt was used to create the basis for this note on generative AI:

"Please Write a note on generative AI to be included in our master's thesis. Clarify that no text or content from generative AI has been used in the forging of the text. The reasons are that we're unable to check every output and, e.g., do not trust prompts such as 'improve the clarity of this paragraph', as we notice that the tool changes words despite being told not to do so. Include this prompt in the appendix, saying only this created output was used in parts."

Source (ChatGPT-4o, August 31st, 2024): https://chatgpt.com/share/1dd28231-3a10-4442-b655-a98275158d1b