Measuring Open Innovation through Textual Analysis: An Assessment of Nonlinear Performance Implications

Thomas Schaeper
University of Muenster, thomas.schaeper@wiwi.uni-muenster.de

Christopher Jung
WWU Muenster, christopher.jung@wiwi.uni-muenster.de

J. Nils Foege
University of Hannover, nils.foege@wa.uni-hannover.de

Stephan Nüesch
University of Muenster, nueesch@wwu.de

Measuring Open Innovation through Textual Analysis: An Assessment of Nonlinear Performance Implications

Completed Research Paper

Thomas Schäper
University of Muenster
Business Management Group
Georgskommende 26
D-48143 Muenster
thomas.schaeper@wiwi.uni-muenster.de

Christopher Jung
University of Muenster
Business Management Group
Georgskommende 26
D-48143 Muenster
christopher.jung@wiwi.uni-muenster.de

J. Nils Foege
University of Hannover
Innovation Management Group
Schloßwender Str. 7
D-30159 Hannover
nils.foege@wa.uni-hannover.de

Stephan Nüesch
University of Muenster
Business Management Group
Georgskommende 26
D-48143 Muenster
stephan.nueesch@wiwi.uni-muenster.de

Abstract

Acknowledging that valuable ideas can come both from inside and outside the organization, firms have turned to more open models of innovation. However, it remains difficult to measure firms' open innovation activities, particularly in a quantitative, longitudinal setting. This impedes an adequate assessment of open innovations' long-term implications for firms' financial performance. While researchers have devoted a considerable effort to examining the link between open innovation and innovative performance, the link to financial performance is unclear. In this study, we develop and validate a text-based measure for firms' open innovation activities, and probe related performance implications in a longitudinal, cross-industry setting. Combining machine-learning content analysis to create an open innovation dictionary, we analyze the 10-K annual reports of 9,100 publicly listed firms in the U.S. between 1994 and 2017. Our results support our theorizing that a nonlinear relationship takes an S-shape between open innovation and financial performance.

Keywords: Open innovation, performance, text-based measure, content analysis

Introduction

The pursuit of sustainable growth and competitive advantage through innovation is one of the key objectives of strategic management. Firms have traditionally dedicated substantial amounts of resources to their internal research and development (R&D) activities (Chesbrough and Crowther 2006). Today's highly dynamic environments, however, have rendered such closed models of innovation increasingly obsolete (Chesbrough 2010). Acknowledging that valuable ideas can come both from inside and outside the organization, firms have turned to more open models of innovation (Chesbrough et al. 2014). Open innovation enables the exchange of ideas, knowledge, and technologies across organizational boundaries, enhancing firms' internal ideation processes, broadening their pool of available knowledge, and, perhaps, reducing the risks and the expenses associated with R&D activities (Terwiesch and Xu 2008). Open innovation has, therefore, become critical for the success of firms' innovation processes (Bogers et al. 2017; Foege et al. 2019; Laursen and Salter 2006; Stanko et al. 2017).

Scholars argue that open innovation can drive both innovative and financial performance (e.g., Faems et al. 2010; Gassmann et al. 2010; Hopkins et al. 2011; Laursen and Salter 2006), but emphasize that it is necessary to empirically assess these open innovation-performance relationships (Du et al. 2014; Faems et al. 2010). However, while there are numerous studies about the effects of open innovation on the outcomes of innovation (e.g., Foege et al. 2017; Grimpe and Kaiser 2010; Laursen and Salter 2006; Salge et al. 2012), the empirical assessment of how it affects performance, especially financial performance, offers fertile ground for scholarly insights (e.g., Bogers et al. 2017; Chesbrough et al. 2014).

The main challenge is quantifying open innovation in a large-scale longitudinal setting, as archival data on firms' open innovation activities is largely unavailable, and firms' internal analyses are highly idiosyncratic to each one's innovation strategy, making it difficult to make inter-firm comparisons (West 2014; West et al. 2014). In this study, we draw on a growing body of literature from various disciplines such as finance, management, and accounting that leverages techniques of textual analysis to quantify organizational behavior (Bellstam et al. 2017; Hoberg and Lewis 2017; Nadkarni and Chen 2014). We employ these techniques to capture open innovation, with the following questions in mind: *How can open innovation be quantified, particularly in large-scale and longitudinal settings? How do open innovation activities influence organizational financial performance?*

To address these questions, we employ a machine-learning approach to create a dictionary on open innovation based on the 1,000 most-cited articles of scholarly research in this domain. Subsequently, we leverage this dictionary to quantify a firm-level open innovation score by analyzing the 10-K annual reports of 41,035 publicly listed firms in the U.S. between 1994 and 2019. After validating this score against other open innovation and innovation measures, we combine the resulting open innovation score with archival accounting data from the CRSP/Compustat merged database to create a panel dataset of 65,089 firm-year observations and probe our hypotheses. Results of our firm fixed-effects regressions support our theorizing of a nonlinear relationship between open innovation and financial performance taking the functional form of an S-shape.

This study contributes to the literature on open innovation in at least three important ways. First, we extend current research by deriving a text-based measure that captures firms' open innovation activities in a large-scale, longitudinal setting. Since our measure is based on 10-K annual reports that are available for all publicly listed firms, it enables researchers to assess open innovation for a broad range of firms across a broad range of industries. Our approach can be applied to text documents other than annual reports, including analyst reports, newspaper articles, and transcripts of quarterly earnings calls. Second, we provide a better understanding of the performance implications of open innovation, notably for firms' financial performance (e.g., Huizingh 2011). We theorize and find that the relationship between open innovation and firm performance is complex, taking the functional form of an S-shape. While previous research yielded inconclusive results about the total relationship between open innovation and financial performance, we suggest that the S-shape we found aligns well with certain findings of that research that are narrower in scope on the costs and benefits of openness. Third, this study can serve as a blueprint for management scholars on how to leverage textual data on a large scale. Our procedure and resulting measure make it possible to go beyond research that uses cross-sectional data (Faems et al. 2010; Laursen and Salter 2006) and examine the effects of open innovation both over time and across firms and industries.

A Text-based Measure of Open Innovation

Overview

In this section, we describe the construction of our measure. To derive a measure of open innovation through textual analysis, we conducted a two-stage content analysis. This is a statistical technique that objectively and systematically identifies specific characteristics of text data (Hoberg and Lewis 2017), searching text for selected words, ideas, and meanings to identify interpretable topics. The technique relies on two main components: a *dictionary* that contains the most frequent keywords of a theme, in our case of open innovation, and an *algorithm* that uses this dictionary to analyze textual information (Goldsmith-Pinkham et al. 2016; Hoberg and Lewis 2017; Lowry et al. 2016).

Open Innovation Dictionary

To create our *open innovation dictionary*, we followed and combined open- and closed-language approaches (Harrison et al. 2019). First, we used Google Scholar to search for the one thousand most-cited studies of open innovation to establish the input text data for creating the dictionary. To capture all relevant studies, we ran two search queries, one searching for the term *open innovation* in the titles of articles and another in their keywords. We then combined the two queries and deleted double entries. We carefully cleaned the sample of texts for all text documents that were unrelated to open innovation per se[1] and excluded studies that were not in English. Furthermore, we cleaned and unified the data through (1) converting all text to lowercase, (2) removing common English stopwords, (3) stemming all words, and (4) using n-gram routines to include sequences of words (Antons et al. 2016; Bodnaruk et al. 2015; Hoberg and Maksimovic 2015). The latter is important to ensure that the identification of sequences of words captures specific concepts that occur as textual compounds (e.g., open innovation, outbound innovation, and knowledge transfer). Our final dataset consists of 925 studies with about 6 million words in total.

Second, we applied a supervised machine learning algorithm based on the Naive Bayes logic to identify proper categories and keywords describing the content of the data (Behl et al. 2014). Machine-based concept identification exhibits close agreement with the judgment of human experts and is, thus, a suitable approach in our context (Campbell et al. 2011). Moreover, it provides reliable and reproducible concept extraction and thematic clustering without human biases (Randhawa et al. 2016). The algorithm identifies key concepts and themes within open innovation literature based on the frequency and co-occurrence of keywords. It also extracts the relationship between these concepts and proposes a name for each category (Randhawa et al. 2016).

To obtain a rigorous open innovation dictionary that describes the concept of open innovation, we applied a two-stage process and experimented with different numbers of topics (Antons et al. 2016; Bellstam et al. 2017). First, we manually computed different models with 15 to 50 topics in steps of 5. We found that models with fewer topics captured the landscape of open innovation similarly well, whereas models with a greater number of topics (> 35 topics) exhibited redundancies and did not combine topics that were substantively close together. We used a model with 23 topics with 323 words as our baseline model for further adaptation, as it described open innovation with its respective key themes well. To better understand the meaning of the words and overarching categories, we manually revisited each category on how the keywords were used in the context of the text data. This helps to increase semantic validity and denomination of categories (Krippendorff 2013). These adjustments comprised the supervised component of our machine learning approach, as we excluded or renamed keywords and reran our topic modeling. Second, we brought in three academic researchers, experts in the field of open innovation, to independently assess the proposed topics and keywords, and, relatedly, to make qualified adjustments. These adjustments included (1) deleting inadequate keywords, (2) combining categories, and (3) labeling category names. Our final open innovation dictionary consists of 51 words across 9 topics. Table 1 provides an overview of our topics and keywords in the open innovation dictionary.[2]

[1] For example, a study that developed new innovation procedures for open heart surgery was misclassified as relevant for open innovation.

[2] As part of our robustness check, we tested different variations of our dictionary including an unabridged version (see section 4.5 Robustness Checks of this manuscript).

Topic Number	Topic Label	Keywords
1	External & Internal Knowledge Flows	External Sources Internal and External Knowledge
2	Strategic Alliances & External Partners	Alliances Partners Joint Collaboration Relationships Cooperation Strategic Alliances External Partners Joint Ventures
3	Open Innovation	Innovation Open Openness Open Innovation Degree of Openness
4	User Innovation	User Innovation Lead User
5	Knowledge Sharing	Exchange Share Collaborative Knowledge Sharing Knowledge Transfer Knowledge Creation External Knowledge
6	Open Source Software	Open Source Open Source Software
7	Inbound & Outbound Open Innovation	Outbound Inbound Inbound Open Innovation Inbound and Outbound Outbound Open Innovation Outbound Activities Open Innovation Activities Inbound Activities Inbound Open Innovation Activities Inbound Innovation Outbound Innovation
8	Adoption & Diffusion	Diffusion Adoption Adoption and Diffusion Diffusion and Adoption Open Innovation Adoption
9	Exploration & Exploitation	Exploitation Exploration Exploration and Exploitation Technology Exploitation Knowledge Exploration Knowledge Exploitation External Technology Exploitation Technology Exploration

Table 1. Topics and Keywords of this Study's Open Innovation Dictionary

Calculating the Open Innovation Score

After deriving the open innovation dictionary, we applied it to textual data. Using the central index key as a company identifier, we used an application programming interface to download all 10-K annual reports for the years 1994 to 2019 from the United States Securities and Exchange Commission (SEC) database. Our obtained data consists of around 245,000 10-K annual reports of 41,035 publicly listed firms in the U.S. between 1994 and 2019. We used our open innovation dictionary to feed an algorithm that analyzed these 10-K annual reports. This algorithm was set up to analyze text data with our customized dictionary and to determine the extent to which the text presented a particular theme. To ensure adequate analysis of the textual data, the algorithm automatically removed common stopwords and blank spaces, and stemmed all words. In determining the final score, we built on prior representations of text-based measures (e.g., Hubbard et al. 2018; Moss et al. 2018; Uotila et al. 2009) and calculated the *open innovation score* ($\theta_{f,t}$) as the number of words that are common to both the text and the open innovation dictionary ($D_{f,t}$) in relation to the length of each annual report ($T_{f,t}$) for each firm f in year t:

(1)
$$\theta_{f,t} = \frac{D_{f,t}}{T_{f,t}}$$

The score lies in the range between 0 and 1, while the empirical observable maximum is at 4.42 percent. Intuitively, this score can be interpreted as a representation of a firm's open innovation activity relative to its overall business activity. Given that we analyzed large scale text data, the algorithm required a lot of computing power. Hence, we used a high-performance computer cluster to calculate our final measure. We employed the largest partition of the cluster with three nodes, 144 CPU cores per node, and 3000GB of memory per node to calculate our final measure. The algorithm reconciled the 10.5 billion words of all 10-K annual reports with the 51 keywords in our open innovation dictionary. This resulted in approximately 535.5 billion instances. The entire calculation took about 65 hours of processing time.

Analyzing the Open Innovation Score

Four potential trends can be identified by analyzing the time series trend of our open innovation measure over the years 1994 to 2017. Beginning from our sample period in 1994, the open innovation score increases and stays relatively stable until the year 2000. We tentatively assume that the bursting of the dot-com bubble in 2000 – which resulted in the failure of many first-generation Internet firms and the subsequent slowdown of the whole economy – dampened activity in open innovation. The next upswing can be observed in 2004, shortly after Henry Chesbrough published his seminal book on open innovation (Chesbrough 2003), which seems to have shifted the attention of many practitioners to the concept (as a measure of the book's ongoing relevance, as of March 2020, it had been cited in academic journals more than 20,000 times). Further, there is a dramatic decrease in the open innovation score after the financial meltdown in the fall of 2008. We tentatively suggest that as firms' funding dried up after the crisis, so did the innovative activity, including in the area of open innovation, which can be costly. As for the recent trend since 2015, the average open innovation score begins to increase while remaining below the levels of the '90s. In sum, this evidence tentatively indicates that open innovation activity varies over time and is shaped by the market environment.

The highest scores of open innovation activity can be observed for electronics, pharmaceutical, and computer firms – those commonly denoted as high-tech (e.g., Hecker 1999). This finding is in line with Chesbrough and Crowther (2006), who underscore not only the importance of open innovation in high-tech but also its growing importance in knowledge-intensive industries. The lowest scores of open innovation activity are found in the energy supply, insurance, and coal industries. This is not surprising, as in these industries the demand is rather stable, and in some cases, prices are regulated (Chaganti and Sambharya 1987; Nason and Patel 2016).

Comparison to other Measures

Following the methodological approach of Demerjian et al. (2012) and Hoberg and Maksimovic (2015), we validate our open innovation score by comparing it to (1) other, less sophisticated proxies of open innovation, and (2) general measures of firms innovation activity that are commonly used in management studies.

We used data on alliances from the SDC Platinum database to validate our open innovation score against firm-level data on firms' alliances. This approach builds on the premise that open innovation activities, particularly those that are outbound, serve as valid proxies for firms' search for innovation outside their boundaries (Dahlander and Gann 2010; Huizingh 2011; West and Bogers 2014). Following this reasoning, we extracted the number of (1) alliances, (2) alliance partners, (3) R&D alliances, and (4) technology transfer alliances as tentative indications of external searches for innovation (e.g., Huang and Rice 2012; Laursen and Salter 2006; Leiponen and Helfat 2010). We, further, compared our measure against common innovation measures, such as (1) R&D intensity, (2) number patent applications, and (3) citation-weighted value of patents (e.g., Blagoeva et al. 2019; Kogan et al. 2016). We present correlations in Table 2.

	Open Innovation Score	*Number of Alliances*	*Number of Alliance Partners*	*Number of R&D Alliances*	*Number of Technology Transfer Alliances*	*R&D Intensity*	*Number of Patent Applications*	*Citation-weighted Value of Patents*
Open Innovation Score		0.08*	0.05*	0.06*	0.03	0.06*	0.05*	0.07*
Number of Alliances	0.05*		0.39*	0.36*	0.33*	0.02*	0.33*	-0.06*
Number of Alliance Partners	0.03*	0.16*		0.23*	0.13*	-0.11*	0.22*	-0.11*
Number of R&D Alliances	0.04*	0.62*	0.10*		0.21*	0.18*	0.25*	0.10*
Number of Technology Transfer Alliances	0.04*	0.46*	0.06*	0.37*		0.18*	0.14*	0.13*
R&D Intensity	0.07*	0.04*	-0.08*	0.11*	0.12*		0.02	0.55*
Number of Patent Applications	0.01*	0.48*	0.07*	0.26*	0.23*	0.04*		0.20*
Citation-weighted Value of Patents	0.03*	-0.04*	-0.06*	-0.02	-0.01	0.26*	$9e^{-4}$	

Notes: R&D Intensity is calculated as the R&D expenditures scaled by sales. Spearman correlations are presented in the upper right, while Pearson correlations are presented in the lower left. Values with a star indicate statistical significance at the 5% level.

Table 2. Comparison with other Measures - Correlations

Overall, these comparisons between our open innovation score and other indications of both open innovation activity and innovation show that the open innovation score exhibits a consistently positive relationship with other innovation scores. While the correlations are quite low, they still provide preliminary indications that our text-based measure captures information on open innovation activity beyond accounting variables. As a benchmark, Hoberg and Lewis (2017) and Bodnaruk et al. (2015) also report low correlations between their text-based measures and accounting data. Our open innovation measure contains relevant information beyond quantitative measures of open innovation, focuses on the novel uses, and offers greater applicability compared to other traditional measures. However, in line with theory, open innovation is a specific form of innovation activity that requires measures that adequately mirror it.

Conceptual Background and Hypotheses

The Nature of Open Innovation

Since the seminal work of Chesbrough (2003), the concept of open innovation has received considerable attention from both academics and managers and has made a substantial impact on research and practice around firm innovation (Randhawa et al. 2016; Stanko et al. 2017). Open innovation is a "distributed innovation process based on purposively managed knowledge flows across organizational boundaries, using pecuniary and non-pecuniary mechanisms in line with the organization's business model" (Chesbrough and Bogers 2014, p. 27). As such, open innovation includes multi-directional knowledge flows and often collaboration with a broad variety of different partners (e.g., Bogers et al. 2017; Foege et al. 2019). Information systems play a critical role in this context, as they provide the network and software for collaboration and idea sharing (Doan et al. 2011). For instance, many firms use information technology systems to implement processes for collaboration in open innovation activities (Bassellier and Benbasat 2004). Moreover, numerous IT-based applications support the development of innovations and control systems for collaborative efforts (Farrell 2003). With the ongoing digitalization, information systems gains in importance for successful open innovation practices (Tarafdar and Gordon 2007).

Given its multi-directional nature, open innovation can be inbound, outbound, or coupled – both inbound and outbound (Enkel and Gassmann 2010). Inbound open innovation refers to firms' sourcing of external knowledge, which includes the acquisition of external expertise in marketplaces as well as the sourcing of external knowledge and ideas. Outbound open innovation is the sale of internal knowledge and technologies and the free revealing of internal resources to external actors (Dahlander and Gann 2010; Foege et al. 2019). Coupled open innovation combines inbound and outbound open innovation (Gassmann et al. 2010; Huizingh 2011; West et al. 2014) and materializes strategic R&D alliances, innovation ecosystems, and innovation collaboration (Stanko et al. 2017).

Open innovation activities often include collaboration with distinct partners including users and customers, suppliers, competitors, scientific organizations, and governmental agencies (e.g., Laursen and Salter 2006; Leiponen and Helfat 2010; Stanko et al. 2017). Scholars suggest and find that the pursuit of open innovation as part of an ecosystem with diverse actors can enhance innovation performance by broadening access to information on technologies and markets, as well as sharing the costs and risks involved in R&D processes (Grimpe and Kaiser 2010; Salge et al. 2012). Research broadly confirms that using heterogeneous types of bidirectional knowledge flows and collaborating with a wide variety of partners can enhance firms' performance (e.g., Dahlander and Gann 2010) by accelerating new product development and reducing time-to-market for new products and services (Faems et al. 2010). Open innovation activities enable firms to capture market share and reap the returns from their innovations (Lauritzen and Karafyllia 2019). Yet a large number of collaboration partners can enhance the complexity of managing the innovation process and can expose the firm to value appropriation resulting from competitors' imitative efforts (Foege et al. 2017; Laursen and Salter 2006; Li et al. 2012). Open innovation can carry the risk of leaking critical knowledge to competitors, who can use this information to develop and market competing technologies, products, and services (Veer et al. 2016) without remunerating the focal firm (Foege et al., 2019). Such opportunistic behavior can lead to significant revenue losses and diminish firms' performance (Li et al. 2012).

Despite these notable examples, studies on the potential benefits and costs of open innovation for financial performance remain scarce. Given the prevalence of open innovation for theory and practice, further studies are warranted to account for the costs and the benefits of open innovation at different levels of openness.

Benefits of Open Innovation

Firms that engage in open innovation draw on a broad variety of knowledge from different fields that are often new to the focal firm (Hoffmann 2005). Collaboration with a broad set of diverse partners of different types (e.g., customers, suppliers, and scientific institutes), who are often globally dispersed (Faems et al. 2010; Foege et al. 2017), enhances firms' possibilities to access valuable knowledge, technologies, and markets (Dahlander and Gann, 2010). Firms that have the required absorptive capacity to successfully receive and process such external knowledge can enhance their opportunities to recombine it with their existing internal knowledge to spark innovation (Cohen and Levinthal 1990; Ritala and Hurmelinna-Laukkanen 2013).

By this means, open innovation can help firms to overcome the myopia of organizational learning, i.e., the tendency to focus on the exploitation of existing knowledge and to neglect the exploration of new knowledge. This myopia is said to be rooted in the disregard of distant places, as well as in the failure to develop new knowledge (Levinthal and March 1993) and to search beyond organizational boundaries (Lopez-Vega et al. 2016).

Firms can create synergies with their partners by jointly leveraging marketing activities and R&D (Chesbrough et al. 2014). At the same time, they can maintain or improve their innovative efforts relative to firms that rely only on internal R&D. Beyond that, open innovation can enable firms to commercialize otherwise unused internal ideas and knowledge outside their boundaries through selling or out-licensing them to their partners, who utilize this information to create new products and sell them to their customers (Chesbrough and Crowther 2006). In sum, open innovation is associated with benefits including accessing valuable external knowledge (Dahlander and Gann 2010), realizing synergies in R&D activities (Chesbrough et al. 2014), and generating additional revenue streams through commercializing unused internal knowledge (Chesbrough and Crowther 2006). All of these benefits can amplify a firm's financial performance (Chesbrough and Rosenbloom 2002).

Costs of Open Innovation

Several types of expenses are associated with open innovation, including setup costs, operating costs, and coordination costs. Setup costs are the costs of transferring existing resources to new areas of operation (Hashai 2015). These one-time costs are exclusive to the start-up of open innovation, as firms have to build new structures for collaboration (Kale and Singh 2010). These structures include the recruitment of new employees or shifting and training of existing ones (Helfat and Eisenhardt 2004), the building of internal management systems, or dedicated business units (Tan and Mahoney 2006), and the purchasing or modification of equipment (Hashai 2015). These initial investments are often sunk. Setup costs will diminish at higher levels of open innovation.

Operating costs are continuous expenses and associated with the maintenance and administration of the newly-created structures (e.g. distribution of staff, travel expenses, rent). These operating costs often ameliorate inefficiencies in the beginning, when existing routines or knowledge are missing. These inefficiencies can ultimately distort operations and lead to an imperfect allocation of resources (Cohen and Levinthal 1990; Fernhaber and Patel 2012; Teece 1980). Over time, operating costs decline steadily as firms build up the required knowledge and routines in the new areas of operation to improve efficiency; for instance, improving the distribution of staff reduces wage costs. Operating costs are likely to decrease with higher levels of openness.

Coordination costs are linked to the management of the partnership portfolio, such as the costs associated with seeking, processing, transferring, and protecting knowledge. They increase with a growing number of open innovation partners and initiatives. Some of them arise from the considerable costs of identifying, assimilating, and utilizing a multitude of external knowledge inputs at the same time (Dahlander and Gann 2010). In this vein, it has been noted that searching too broadly may lead to a set of ideas and opportunities that (1) is too large to be effectively managed, (2) comes at the wrong time and place, and (3) cannot be sufficiently pursued (Katila and Ahuja 2002; Koput 1997). Another important type of coordination costs in collaborative research projects arises from firms' protective efforts (e.g. patents, copyrights, trademarks) and litigation cases to claim ownership over ambiguous intellectual property (Foege et al., 2017, Foege et al., 2019). The costs to coordinate and deal with these legal issues should not be underestimated (Li, Eden, Hitt, Ireland, and Garrett, 2012). Coordination costs are to likely increase over time, especially as the intensity of engaging in open innovation activities increases.

Financial Performance at Various Levels of Open Innovation

We now synthesize how the benefits and costs mentioned above vary at different levels of open innovation activity. Figure 1 visualizes our theorizing.

Performance at Low Levels of Open Innovation

At low levels of open innovation activity, the expected benefits of open innovation are modest, as the possibilities for realizing synergies and exploiting knowledge are limited. While coordination costs are negligible at low levels of open innovation activity, as the number of partners is manageable with low transaction costs, setup costs can be substantial, as the structures for collaboration and joint activities must be established. Moreover, at low levels of open innovation, firms lack supporting routines and knowledge (Fernhaber and Patel 2012), making it easier to err in business decisions and making it harder to allocate resources effectively. In sum, we expect that high setup and operating costs undermine the initial benefits of open innovation at low levels of openness. This is visualized in the left part of Figure 1. We, therefore, hypothesize:

Hypothesis 1. At low levels of open innovation, the relationship between open innovation and firm performance is negative.

Performance at Moderate Levels of Open Innovation

The benefits that firms can draw from open innovation are likely to increase the more they engage in it. Eventually, they can recuperate the setup costs and operating costs. With a growing number of partners and collaboration projects, they can realize synergies in R&D, build up experience in combining and absorbing external knowledge, and commercialize unused knowledge (Dahlander and Gann 2010; Lauritzen and Karafyllia 2019). While the operating costs can be expected to decline at moderate levels of open innovation, the coordination costs will increase, though only slightly as long as the number of partners remains fairly manageable. Operating costs decrease as managers can use their built-up knowledge and experience to improve the resource allocations, avoid mistakes in business decisions, and apply more nuanced management of operations. At this point, routines become established, helping to reduce inefficiencies, and, therefore, minimize operating costs.

Taken together, we expect that at moderate levels of open innovation activities, the benefits of open innovation outweigh the negatives: the operating costs are low and the coordination costs only moderate. This relationship is visualized in the center of Figure 1. We, therefore, hypothesize:

Hypothesis 2. At moderate levels of open innovation, the relationship between open innovation and firm performance is positive.

Performance at High Levels of Open Innovation

Despite its immense benefits, engaging in open innovation is also subject to diminishing returns. At high levels of open innovation, the beneficial synergies reach a plateau, while the coordination costs skyrocket as the costs of searching for knowledge and partners increase exponentially and deplete firm resources (Bogers et al. 2017; Huizingh 2011; Salge et al. 2012; Stanko et al. 2017). For example, the transaction costs for transferring technologies among more than two firms become exponentially higher than between only two firms as the contractual arrangements become more complex (Kale et al. 2000). The larger the number of partners a firm collaborates with, the greater the complexity of managing this portfolio of partnerships, while the returns diminish due to searching and coordination costs (Foege et al., 2017). The complexity can seriously dampen the efficiency of R&D investments among a larger set of partners (Dahlander and Gann 2010), such that we expect an increasingly negative effect of coordination costs on firm performance when extending open innovation efforts beyond a certain threshold. The more firms engage in open innovation, the greater the threat of opportunistic behavior looms (Laursen and Salter 2014).

In sum, we expect that the benefits of open innovation to be outweighed by increasing coordination costs and opportunistic behavior by other firms at high levels of open innovation. This is visualized in the right part of Figure 1. We, therefore, suggest:

Hypothesis 3. At high levels of open innovation, the relationship between open innovation and firm performance is negative.

All things considered, after establishing viable structures and related operations, firms can benefit from opening their innovation activities. However, the payoff of open innovation activities reaches a turning point, after which negative performance effects emerge. Based on our hypotheses, as open innovation begins and increases, setup costs and operating costs lead to benefits, but then coordination costs take over, resulting in an S-shape. Figure 1 portrays this nonlinear S-shaped relationship between open innovation and financial performance, with the slope being negative at low levels of open innovation, positive at moderate levels of open innovation, and negative again at high levels of open innovation.

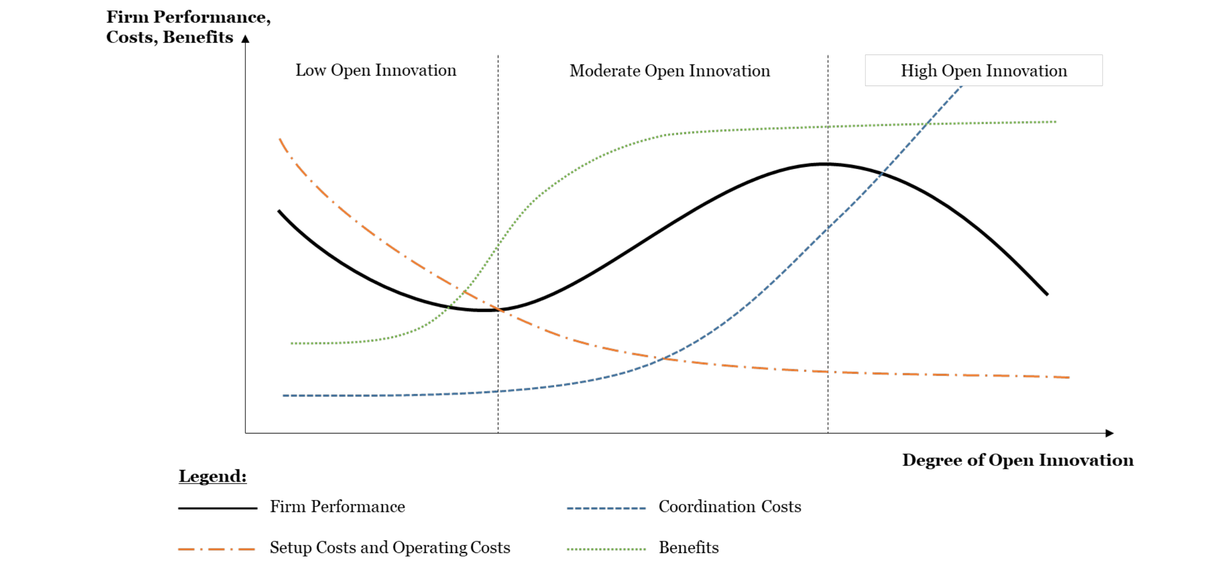


Figure 1. The Theoretical Open Innovation–Firm Performance Relationship

Empirical Methodology

In this section, we describe the empirical methodology we employed. In doing so, we explain the data construction and variable selection of our panel dataset, and we develop the main empirical analysis. We utilize our text-based measure of open innovation to examine related performance implications and probe our hypotheses.

Sample Construction

To probe the hypothesized S-shaped relationship between open innovation and financial performance, we merged our open innovation score with data on security prices and accounting data from the CRSP/Compustat merged database (CCM), including controls of common drivers of firm performance (e.g., Deb et al. 2017; Kim and Bettis 2014). To achieve consistency with prior research, we excluded firms that belonged to the idiosyncratic industries of utilities (SIC 4900-4999), financial institutions (SIC 6000-6999), governmental organizations (SIC 9100-9199), and non-classifiable establishments (SIC 9900-9999). Our final baseline sample comprised 65,089 firm-year observations of publicly listed firms in the U.S. between 1994 and 2017. To rule out the possibility of outliers distorting our inferences, we winsorized the data on the 1st and 99th percentile (e.g., Deb et al. 2017; Haans et al. 2016).

Variable Definitions

Dependent Variable. We used firms' Total *Q,* an updated variant of Tobin's *q,* to capture firms' performance. Total *Q* is defined as the ratio of (1) the sum of the market value of outstanding equity and the book value of outstanding debt less the current assets of a firm in the numerator, and (2) the book value of physical capital and intangible capital in the denominator (Peters and Taylor 2017). As intangible assets become increasingly important, with intangible capital accounting up to 34% of a firm's total capital (Corrado and Hulten 2010), and industries shifting towards becoming more service- and technology-based, incorporating intangible capital – which is disregarded by Tobin's *q* – is indispensable (Peters and Taylor, 2017). The intangible assets are measured as the sum of a firm's (1) knowledge and (2) organizational capital using the perpetual inventory method; their inclusion provides a rigorous estimate of the firm's intangible resources.[3] This study aimed to assess the performance implications of open innovation activities, and the use of a market-based measure of firm performance has been recommended in prior studies related to innovation (e.g., Bharadwaj et al. 1999; Fosfuri and Giarratana 2009); further, it aligns well with our theorizing. Total *Q* is an approximation of a firm's long-term profitability, and its underlying

[3] Total *Q* is available via the Peters and Taylor data library (on WRDS). For further details on the calculation of these additional components, please see Peters and Taylor (2017).

approach is common to approximate corporate growth prospects (Erickson and Whited 2012). In extensive robustness checks, Erickson and Whited (2012) document that Total Q outperforms Tobin's q and other measures of growth opportunities.

Explanatory Variables. We used the *open innovation score* as our main independent variable. To account for the non-linearity as part of our hypothesizing, we followed the common approach of probing S-shaped relationships (e.g., Berry and Kaul 2016; Chen et al. 2012; Hashai 2015) and included the second (*open innovation score squared)* and third polynomial terms (*open innovation score cubic)* of the independent variable in our model.

Control Variables. To control for other potential explanations of firm performance, we included several control variables. First, we included *industry Total Q* as the mean value of the dependent variable (four-digit SIC level) to our model. Further, we included *firm size*, defined as the natural logarithm of net sales (Croci and Petmezas 2015), as larger firms are commonly associated with lower growth prospects (Josefy et al. 2015). Similarly, we included *firm growth* as the annual sales growth rate calculated as the natural logarithm of sales growth (Deb et al. 2017; Kim and Bettis 2014). Beyond that, we controlled for the *undistributed cash flow*, a firm's prior potential to generate cash that can trigger firm growth, calculated as the operating income before depreciation, minus the total income taxes, minus the year-to-year changes in deferred taxes, minus the gross interest expenses on total debt, minus the sum of preferred dividend payable on cumulative preferred stock and dividend paid on noncumulative preferred stock, minus the total dollar amount of dividend declared on common stock, and lastly scaled by total assets (Kim and Bettis 2014).

Prior research (e.g., Hall 1992) suggests that market measures of firm value are strongly correlated with research and advertising expenditures. Therefore, we included *R&D intensity*, measured as R&D expenditures divided by net sales, and *advertising intensity*, measured as advertising expenditures divided by net sales (Deb et al. 2017). We replaced missing values of R&D and advertising expenditures with zero, and included two dummies for missing data on R&D and advertising spending respectively (e.g., Blagoeva et al. 2019), which is common practice in research to circumvent potential bias towards companies with R&D-intensive firms (Himmelberg et al. 1999). We capped the upper limits of both intensities at one (Deb et al. 2017; Kim and Bettis 2014).

In addition, we controlled for *potential slack*, defined as the ratio of total debts to total assets. Additional financial resources that can be adopted from the outside of the firm can be used to finance investments and thus facilitate firm growth. The ratio was subtracted from one to facilitate interpretability so that higher values denote high potential slack (Deb et al. 2017). Negative values were dropped to rule out bias from highly indebted firms (Alti 2006). Following Deb et al. (2017) and Kim and Bettis (2014), we included *capital intensity*, calculated as capital expenditures divided by total assets, as higher expenditures are a natural trigger of firm performance. To account for unobserved heterogeneity across time periods, we included *year dummies* (Deb et al. 2017; Kim and Bettis 2014).

Statistical Model

We took several steps to ensure the adequacy of our model specification. First, a Hausman test confirmed the predominance of a fixed-effects model over random-effects. We, therefore, estimated a firm fixed-effects model to prevent unobserved heterogeneity from multiple observations per firm. Second, we detected first-order autocorrelation using a Wooldridge test (Wooldridge 2013) and, as stated before, introduced an industry-adjusted measure of the dependent variable as a control variable (Gentry and Shen 2013). Third, we used a Breusch-Pagan test to find heteroscedasticity; thus, we included robust standard errors clustered at the firm level. Fourth, an analysis of the variance inflation factors showed that multicollinearity was of minor concern, as all values were well below critical thresholds. Lastly, to control for year effects, we included year dummies in the analysis.

Regression Results of the Open Innovation-Firm Performance Relationship

To test our suggested S-shaped relationship between open innovation and firm performance, we estimated the following model:

$$TQ_{f,t} = \beta_0 + \beta_1\theta_{f,t} + \beta_2\theta_{f,t}^2 + \beta_3\theta_{f,t}^3 + \varepsilon_t + \gamma_f + \omega_{f,t} \tag{2}$$

where the dependent variable $TQ_{f,t}$ is Total Q for firm f in year t. The independent variable $\theta_{f,t}$ (and its second polynomial $\theta_{f,t}^2$ and third polynomial $\theta_{f,t}^3$) denotes the open innovation score for firm f in year t.

The model also includes the intercept (β_0), year fixed effects (ε_t), firm fixed effects (γ_f), and control variables ($\omega_{f,t}$). To probe an S-shape functional form, the β_3 coefficient is of key concern. A significant and positive β_3 is associated with a strictly monotonically increasing curve, whereas a significant and negative β_3 is associated with a strictly monotonically decreasing curve.

For the open innovation-firm performance relationship to follow an S-shape functional form, three conditions are to be met (Lind and Mehlum 2010): First, the coefficient β_3 needs to be negative and statistically significant. Second, the slopes before, between, and after the two saddle points, i.e., local maximum and local minimum, need to be sufficiently steep and statistically significant, and negative for the left part, positive for the middle part, and negative again for the right part of the curve. Third, the inflection point of the curve needs to be located within the data range as otherwise the S-shape curve may be incomplete (Haans et al. 2016).

The results are shown in Table 3. As expected, we observed positive effects of several growth-oriented control variables, including industry Total *Q*, potential slack, and capital intensity in Model 1. Model 2 indicates the results of estimating equation (2). We found a negative and statistically significant effect of the *open innovation score* $\theta_{f,t}$ (β = -0.31; p = 0.000), a positive and statistically significant effect of *open innovation score squared* $\theta_{f,t}^2$ (β = 0.16; p = 0.000), and a negative and statistically significant effect of *open innovation score cubic* $\theta_{f,t}^3$ (β = -0.02; p = 0.000) on firm performance. As for the latter, the first condition of an S-shape was fulfilled. To examine the two remaining conditions for an S-shape, further analysis of the first, second, and third derivative of the functional form was necessary.

To probe the remaining conditions, we determined the minimum and maximum, located at point (1.27/0.52) and point (4.06/0.74). For these points, we selected three specific values and inserted them in the first derivation to examine the slopes and to conduct slope tests. We chose (1) the median value between the intercept and the minimum point, (2) the median value between the minimum and maximum point, and (3) the median value between the maximum point and the root of the function. In line with our theorization, we found a negative and significant slope for the left part, a positive and significant slope for the middle part, and a negative and significant slope for the right part of the function, which indicated the existence of an S-shaped relationship between open innovation and firm performance. Further, we split the data based on the two saddle points of our cubic explanatory variable and ran three linear regressions to check if the slopes were consistent with the predicted shape of the curve (Haans et al. 2016). We found that the regression below the first saddle point revealed a negative open innovation-firm performance relationship, the regression below the second saddle point exhibited a positive open innovation-firm performance relationship, and the regression above the second saddle point showed a negative open innovation-firm performance relationship. These findings supported the existence of an S-shaped relationship. Finally, we needed to find the roots of the second derivative to determine the inflection point. Following Haans et al. (2016) and Hirschberg and Lye (2005), we applied the Fieller method and found that the 95 percent confidence interval of the left to right inflection point was located within our data range.[4] In sum, all conditions for an S-shaped relationship were met. The visualization in Figure 2 also confirms an S-shaped relationship between open innovation and firm performance. These findings support our three hypotheses that together suggest an S-shaped functional form for the open innovation-performance relationship.

Robustness Checks

We conducted several checks to show the robustness of our results. First, we calculated our open innovation measure with three dictionaries with varying numbers of topics, and, thus, varying numbers of keywords, and re-examined our main analysis. We used (1) our baseline open innovation dictionary before the expert assessment with 23 topics and 353 words, (2) a reduced dictionary with 14 topics and 134 words, and (3) a more-reduced dictionary with 11 topics and 75 words. In all three cases, our results remain fully robust and confirm an S-shaped relationship between open innovation and financial performance. We, further, extracted the "Business Description" and "Management Discussion and Analysis" section from all 10-K filings and reran our analysis to rule out possible bias from particular parts of the 10-K filing structure. Our

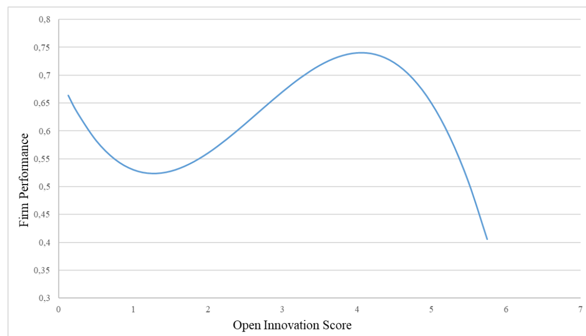
[4] We calculated two Fieller intervals for the left and right part of the function. The range of the 95% confidence interval of the left part is [0.82, 1.37] and [3.40, 4.22] for the right part of the function.

analysis with the extracted parts confirmed the S-shaped relationship in both sections. Second, we conducted a manual coding of 50 10-K filings to control for possible biases from misclassifications of the algorithm. We find a highly significant and positive correlation between our manual coding and the algorithm's results (ß = 0.85, p = 0.000), which shows the good accuracy of our algorithm. Third, we used alternative measures to quantify firm performance. Following Kim and Bettis (2014), we reran our analysis with Tobin's *q* as our dependent variable, defined as the market value of the firm divided by its total assets (Kim and Bettis 2014; O'Brien and Folta 2009). Following common procedures, we dropped all observations of Tobin's *q* that exceeded a threshold of 10 (Kim and Bettis 2014). Our results were fully robust when using Tobin's *q*. Fourth, we included several other control variables that are alternative explanations of firm performance. We introduced a lagged measure of our Total *Q* (Deb et al. 2017), firm age (Demerjian et al. 2012), and amount spent in acquisitions (Fresard 2010) as control variables, and found fully robust results. Fifth, to rule out the influence of sector-level heterogeneity on investment, we confined our sample to the manufacturing industries (NAICS 310000-339999) (Chen 2008; Kim and Kung 2011; Tong et al. 2008), and found, again, that our results were fully robust.

	Model (1)	**Model (2)**
Dependent Variable	Total *Q*	Total *Q*
Controls		
Industry Total *Q*	0.29***	0.30***
	(0.014)	(0.013)
Firm Size	-0.07**	-0.06**
	(0.018)	(0.018)
Firm Growth	0.03***	0.03***
	(0.005)	(0.005)
Undistributed Cashflow	0.45***	0.44***
	(0.035)	(0.035)
R&D Intensity	-0.88***	-0.89***
	(0.190)	(0.191)
Missing R&D Expenditures	0.02	0.01
	(0.041)	(0.042)
Advertising Intensity	-0.63	-0.64
	(0.690)	(0.684)
Missing Advertising Expenditures	0.02	0.03
	(0.032)	(0.033)
Potential Slack	0.54***	0.56***
	(0.060)	(0.061)
Capital Intensity	1.53***	1.50***
	(0.146)	(0.149)
Explanatory		
Open Innovation Score		-0.31***
		(0.079)
Open Innovation Score Squared		0.16***
		(0.038)
Open Innovation Score Cubic		-0.02***
		(0.005)
Intercept	0.60***	0.70***
	(0.120)	(0.132)
R^2	0.14	0.14
Firm-fixed effects	Yes	Yes
Year-fixed effects	Yes	Yes
N	67.281	65.079

Notes: Robust standard errors are reported in parentheses. Statistical significance is reported as *** p < 0.01; ** p < 0.05; * p < 0.1.

Table 3. Fixed-effects Regression Analyses explaining the Effects of Open Innovation on Firm Performance



Note: This graph was constructed using the parameters of the fixed-effects model.

Figure 2. Empirical S-Shaped Relationship between Open Innovation and Firm Performance

Conclusion

In this study, we set out to develop a measure of open innovation, which is a pressing challenge for management scholars (e.g., Bogers et al. 2017; Stanko et al. 2017). To do so, we used machine-learning algorithms to derive a text-based, quantitative measure of open innovation. Building on the one thousand most-cited articles about open innovation, we used advanced topic modeling techniques to create an open innovation dictionary, which we subsequently employed to analyze the 10-K annual reports of 9,100 publicly listed firms in the U.S. between 1995 and 2017. We validated the resulting measure of open innovation by comparing it to other existing proxies of open innovation activity and innovation activities in general to show the adequacy of our new measure. As part of our hypotheses testing, we then merged our measure with archival market and accounting data. Our results suggest that the relationship between open innovation and financial performance follows an S-shape. Such a relationship confirms our theorizing that engaging in open innovation decreases performance initially, then increases performance, but with diminishing returns, so that at high levels of open innovation activity, the performance contribution decreases. Initial set-up costs and ongoing operating costs alongside exponentially increasing costs of search, idea selection, and coordination (Dahlander and Gann 2010; Foege et al. 2019; Veer et al. 2016) serve as theoretical explanations for the S-shaped relationship.

This study contributes to the scholarly literature on open innovation, providing a text-based measure that offers an updated way to gauge firms' open innovation activity. Building on our open innovation dictionary, our machine-learning algorithm is versatile, as it can assess firms' open innovation activities using any form of text input including analysts' reports, product announcements, and even newspaper articles. This is especially useful for assessing open innovation activities of smaller, non-listed firms that are not required to publish annual reports.

Our open innovation measure is based on publicly available annual reports to capture firms' open innovation activities in a cross-industry, longitudinal setting. This approach circumvents the shortcomings of using corporate surveys, which are often limited to a cross-sectional setting. By this means, our measure enables open innovation scholars to leverage the concept of open innovation in new ways, particularly by combining it with large archival datasets that include information about firms' attributes and various performance measures (e.g., Bogers et al. 2017; Huizingh 2011). This can enable future research to relate open innovation to firm-internal factors such as resources (e.g., Barney 1991), managerial capabilities (Sirmon et al. 2011), top management teams (Hambrick and Mason 1984), and organizational structure (DeCanio et al. 2000), as well as to firm-external factors such as the organizational task environment (Dess and Beard 1984), institutional arrangements (Nelson and Nelson 2002), and social capital (Tsai and Ghoshal 1998) – to name a few examples. This is of particular theoretical and empirical interest, as this study enables researchers to connect open innovation with major management theories such as the resource-based view (Barney 1991), the behavioral theory of the firm (Cyert and March 1963), and the dynamic capabilities view (Teece et al. 1997), and empirically probe the resulting models.

Beyond that, our study shows how different degrees of openness relate to firms' performance and, therefore, provides more nuanced theorizing of open innovation at varying levels of activity and firm performance. Our results on the open innovation-firm performance relationship show that it follows an S-shape. This challenges the presumed inverted U-shaped relationship between open innovation and innovation performance (e.g., Grimpe and Kaiser 2010; Laursen and Salter 2006; Leiponen and Helfat 2010; Salge et al. 2012). In fact, the open innovation-performance relationship seems to be more complex, as an inverted U-shaped relationship can be found only between intermediate and high levels of open innovation activity. From low to moderate levels, the relationship is best described as a U-shape – which is part of the S-shape documented in the present study.

Our study has some limitations that offer fruitful ground for future research. First, as for our content analysis, it is important to determine an appropriate list of words and a dictionary describing the open innovation landscape. While our algorithm processed and categorized frequent words, three experts revised the list and excluded inappropriate words, which can have led to subjective bias. However, our cautious and conservative approach ensured only modest levels of subjectivity, if at all. Second, the data used in this study is related to publicly listed firms from the U.S. Thus, this study focuses on larger companies located in only one of the many global economies. It would hence be interesting to analyze small and medium-sized companies and companies from other geographical areas to check if our findings hold for these cases as well. Third, we validate our findings for different dependent variables, comprising measures of financial performance. However, activities of open innovation can have further consequential features. Most notably, such features comprise effects on the real economy, such as disrupting a nascent industry, eroding the distribution of market shares, or triggering technological dynamism.

References

Alti, A. 2006. "How Persistent Is the Impact of Market Timing on Capital Structure?" *The Journal of Finance* (61:4), pp. 1681-1710.

Antons, D., Kleer, R., and Salge, T. O. 2016. "Mapping the Topic Landscape of JPIM, 1984-2013: In Search of Hidden Structures and Development Trajectories," *Journal of Product Innovation Management* (33:6), pp. 726-749.

Barney, J. 1991. “Firm Resources and Sustained Competitive Advantage,” *Journal of Management* (17:1), pp. 99-120.

Bassellier, G., Benbasat, I. 2004. "Business Competence of Information Technology Professionals: Conceptual Development and Influence on IT-business Partnerships," *MIS Quarterly* (28:4), pp. 394–673.

Behl, D., Handa, S., and Arora, A. 2014. “A Bug Mining Tool to Identify and Analyze Security Bugs using Naive Bayes and TF-IDF,” *2014 International Conference on Reliability Optimization and Information Technology (ICROIT)*, pp. 294-299.

Bellstam, G., Bhagat, S., and Cookson, J. A. 2017. “A Text-Based Analysis of Corporate Innovation,” *SSRN Electronic Journal*.

Berry, H., and Kaul, A. 2016. “Replicating the Multinationality-Performance Relationship: Is there an S-curve?” *Strategic Management Journal* (37:11), pp. 2275-2290.

Bharadwaj, A. S., Bharadwaj, S. G., and Konsynski, B. R. 1999. “Information Technology Effects on Firm Performance as Measured by Tobin's q,” *Management Science* (45:7), pp. 1008-1024.

Blagoeva, R., Mom, T. J.M., Jansen, J. J.P., and George, G. 2019. “Problem-Solving or Self-Enhancement? A Power Perspective on How CEOs Affect R&D Search in the Face of Inconsistent Feedback,” *Academy of Management Journal*.

Bodnaruk, A., Loughran, T., and McDonald, B. 2015. “Using 10-K Text to Gauge Financial Constraints,” *Journal of Financial and Quantitative Analysis* (50:4), pp. 623-646.

Bogers, M., Zobel, A.-K., Afuah, A., Almirall, E., Brunswicker, S., Dahlander, L., Frederiksen, L., Gawer, A., Gruber, M., Haefliger, S., Hagedoorn, J., Hilgers, D., Laursen, K., Magnusson, M. G., Majchrzak, A., McCarthy, I. P., Moeslein, K. M., Nambisan, S., Piller, F. T., Radziwon, A., Rossi-Lamastra, C., Sims, J., and Ter Wal, A. L. J. 2017. “The Open Innovation Research Landscape: Established Perspectives and Emerging Themes across different Levels of Analysis,” *Industry and Innovation* (24:1), pp. 8-40.

Campbell, C., Pitt, L. F., Parent, M., and Berthon, P. R. 2011. “Understanding Consumer Conversations Around Ads in a Web 2.0 World,” *Journal of Advertising* (40:1), pp. 87-102.

Chaganti, R., and Sambharya, R. 1987. “Strategic Orientation and Characteristics of Upper Management,” *Strategic Management Journal* (8:4), pp. 393-401.

Chen, C.-J., Huang, Y.-F., and Lin, B.-W. 2012. “How Firms Innovate through R&D Internationalization? An S-curve Hypothesis,” *Research Policy* (41:9), pp. 1544-1554.

Chen, W.-R. 2008. “Determinants of Firms' Backward- and Forward-Looking R&D Search Behavior,” *Organization Science* (19:4), pp. 609-622.

Chesbrough, H., and Bogers, M. 2014. “Explicating Open Innovation: Clarifying an Emerging Paradigm for Understanding Innovation,” in *New Frontiers in Open Innovation*, H. Chesbrough, W. Vanhaverbeke and J. West (eds.), Oxford University Press, pp. 3-28.

Chesbrough, H., and Crowther, A. K. 2006. “Beyond High Tech: Early Adopters of Open Innovation in other Industries,” *R and D Management* (36:3), pp. 229-236.

Chesbrough, H., and Rosenbloom, R. 2002. “The Role of the Business Model in Capturing Value from Innovation: Evidence from Xerox Corporation's Technology Spin-off Companies,” *Industrial and Corporate Change* (11:3), pp. 529-555.

Chesbrough, H., Vanhaverbeke, W., and West, J. (eds.) 2014. *New Frontiers in Open Innovation*, Oxford University Press.

Chesbrough, H. W. 2003. *Open innovation*: *The New Imperative for Creating and Profiting from Technology*, Boston, Mass.: Harvard Business School Press.

Chesbrough, H. W. 2010. *Open innovation*: *The New Imperative for Ceating and Profiting from Technology*, Boston, Mass.: Harvard Business School Press.

Cohen, W. M., and Levinthal, D. A. 1990. “Absorptive Capacity: A New Perspective on Learning and Innovation,” *Administrative Science Quarterly* (35:1), pp. 128-152.

Corrado, C. A., and Hulten, C. R. 2010. “How Do You Measure a “Technological Revolution”?” *American Economic Review* (100:2), pp. 99-104.

Croci, E., and Petmezas, D. 2015. “Do Risk-taking Incentives Induce CEOs to Invest? Evidence from Acquisitions,” *Journal of Corporate Finance* (32), pp. 1-23.

Cyert, R. M., and March, J. G. 1963. *A Behavioral Theory of the Firm*, Malden, Mass.: Blackwell.

Dahlander, L., and Gann, D. M. 2010. “How Open Is Innovation?” *Research Policy* (39:6), pp. 699-709.

Deb, P., David, P., and O'Brien, J. 2017. “When Is Cash Good or Bad for Firm Performance?” *Strategic Management Journal* (38:2), pp. 436-454.

DeCanio, S. J., Dibble, C., and Amir-Atefi, K. 2000. "The Importance of Organizational Structure for the Adoption of Innovations," *Management Science* (46:10), pp. 1285-1299.

Demerjian, P., Lev, B., and McVay, S. 2012. "Quantifying Managerial Ability: A New Measure and Validity Tests," *Management Science* (58:7), pp. 1229-1248.

Dess, G. G., and Beard, D. W. 1984. "Dimensions of Organizational Task Environments," *Administrative Science Quarterly* (29:1), pp. 57-73.

Doan, A., Ramakrishnan, R., and Halevy, A.Y. 2011. "Crowdsourcing Systems on the World-wide Web," *Communications of the ACM (*54:4), pp. 87–96.

Du, J., Leten, B., and Vanhaverbeke, W. 2014. "Managing Open Innovation Projects with Science-based and Market-based Partners," *Research Policy* (43:5), pp. 828-840.

Enkel, E., and Gassmann, O. 2010. "Creative Imitation: Exploring the Case of Cross-industry Innovation," *R and D Management* (40:3), pp. 256-270.

Erickson, T., and Whited, T. M. 2012. "Treating Measurement Error in Tobin's q," *Review of Financial Studies* (25:4), pp. 1286-1329.

Faems, D., Visser, M. de, Andries, P., and van Looy, B. 2010. "Technology Alliance Portfolios and Financial Performance: Value-Enhancing and Cost-Increasing Effects of Open Innovation*," *Journal of Product Innovation Management* (27:6), pp. 785-796.

Farrell, D. 2003. "The Real New Economy," *Harvard Business Review*, pp. 2–10.

Fernhaber, S. A., and Patel, P. C. 2012. "How Do Young Firms Manage Product Portfolio Complexity? The Role of Absorptive Capacity and Ambidexterity," *Strategic Management Journal* (33:13), pp. 1516-1539.

Foege, J. N., Lauritzen, G. D., Tietze, F., and Salge, T. O. 2019. "Reconceptualizing the Paradox of Openness: How Solvers Navigate Sharing-protecting Tensions in Crowdsourcing," *Research Policy* (48:6), pp. 1323-1339.

Foege, J. N., Piening, E.P.K., and Salge, T. O. 2017. "Don't Get Caught on the Wrong Foot: A Resource-based Perspective on Imitation Threats in Innovation," *International Journal of Innovation Management* (21:03), p. 1750023.

Fosfuri, A., and Giarratana, M. S. 2009. "Masters of War: Rivals' Product Innovation and New Advertising in Mature Product Markets," *Management Science* (55:2), pp. 181-191.

Fresard, L. 2010. "Financial Strength and Product Market Behavior: The Real Effects of Corporate Cash Holdings," *The Journal of Finance* (65:3), pp. 1097-1122.

Gassmann, O., Enkel, E., and Chesbrough, H. 2010. "The Future of Open Innovation," *R and D Management* (40:3), pp. 213-221.

Gentry, R. J., and Shen, W. 2013. "The Impacts of Performance Relative to Analyst Forecasts and Analyst Coverage on Firm R&D Intensity," *Strategic Management Journal* (34:1), pp. 121-130.

Goldsmith-Pinkham, P., Hirtle, B. J., and Lucca, D. O. 2016. *Parsing the Content of Bank Supervision*, New York, NY: Federal Reserve Bank of New York.

Grimpe, C., and Kaiser, U. 2010. "Balancing Internal and External Knowledge Acquisition: The Gains and Pains from R&D Outsourcing," *Journal of Management Studies* (47:8), pp. 1483-1509.

Haans, R. F. J., Pieters, C., and He, Z.-L. 2016. "Thinking about U: Theorizing and Testing U- and Inverted U-shaped Relationships in Strategy Research," *Strategic Management Journal* (37:7), pp. 1177-1195.

Hall, R. 1992. "The Strategic Analysis of Intangible Resources," *Strategic Management Journal* (13:2), pp. 135-144.

Hambrick, D. C., and Mason, P. A. 1984. "Upper Echelons: The Organization as a Reflection of Its Top Managers," *The Academy of Management Review* (9:2), p. 193.

Harrison, J. S., Thurgood, G. R., Boivie, S., and Pfarrer, M. D. 2019. "Measuring CEO Personality: Developing, Validating, and Testing a Linguistic Tool," *Strategic Management Journal* (40:8), pp. 1316-1330.

Hashai, N. 2015. "Within-industry Diversification and Firm Performance-an S-shaped Hypothesis," *Strategic Management Journal* (36:9), pp. 1378-1400.

Hecker, D. 1999. "High-Technology Employment: A Broader View," *Monthly Labor Review* (122:6).

Helfat, C. E., and Eisenhardt, K. M. 2004. "Inter-temporal Economies of Scope, Organizational Modularity, and the Dynamics of Diversification," *Strategic Management Journal* (25:13), pp. 1217-1232.

Himmelberg, C. P., Hubbard, R.G., and Palia, D. 1999. "Understanding the Determinants of Managerial Ownership and the Link between Ownership and Performance," *Journal of Financial Economics* (53:3), pp. 353-384.

Hirschberg, J. G., and Lye, J. N. 2005. "Inferences for the Extremum of Quadratic Regression Models," *SSRN Electronic Journal*.

Hoberg, G., and Lewis, C. 2017. "Do Fraudulent Firms produce Abnormal Disclosure?" *Journal of Corporate Finance* (43), pp. 58-85.

Hoberg, G., and Maksimovic, V. 2015. "Redefining Financial Constraints: A Text-Based Analysis," *Review of Financial Studies* (28:5), pp. 1312-1352.

Hoffmann, W. H. 2005. "How to Manage a Portfolio of Alliances," *Long Range Planning* (38:2), pp. 121-143.

Hopkins, M. M., Tidd, J., Nightingale, P., and Miller, R. 2011. "Generative and Degenerative Interactions: Positive and Negative Dynamics of Open, User-centric Innovation in Technology and Engineering Consultancies," *R and D Management* (41:1), pp. 44-60.

Huang, F., and Rice, J. 2012. "Openness in Product and Process Innovation," *International Journal of Innovation Management* (16:04), p. 1250020.

Hubbard, T. D., Pollock, T. G., Pfarrer, M. D., and Rindova, V. P. 2018. "Safe Bets or Hot Hands? How Status and Celebrity Influence Strategic Alliance Formations by Newly Public Firms," *Academy of Management Journal* (61:5), pp. 1976-1999.

Huizingh, E. K.R.E. 2011. "Open Innovation: State of the Art and Future Perspectives," *Technovation* (31:1), pp. 2-9.

Josefy, M., Kuban, S., Ireland, R. D., and Hitt, M. A. 2015. "All Things Great and Small: Organizational Size, Boundaries of the Firm, and a Changing Environment," *The Academy of Management Annals* (9:1), pp. 715-802.

Kale, P., and Singh, H. 2010. "Managing Strategic Alliances: What Do We Know Now, and Where Do We Go From Here?" *Strategic Direction* (26:2).

Kale, P., Singh, H., and Perlmutter, H. 2000. "Learning and Protection of Proprietary Assets in Strategic Alliances: Building Relational Capital," *Strategic Management Journal* (21:3), pp. 217-237.

Katila, R., and Ahuja, G. 2002. "Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction," *Academy of Management Journal* (45:6), pp. 1183-1194.

Kim, C., and Bettis, R. A. 2014. "Cash Is Surprisingly Valuable as a Strategic Asset," *Strategic Management Journal* (35:13), pp. 2053-2063.

Kim, H., and Kung, H. 2011. "Asset Specificity, Economic Uncertainty, and Corporate Investment," *SSRN Electronic Journal*.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. 2016. "Technological Innovation, Resource Allocation, and Growth," *SSRN Electronic Journal*.

Koput, K. W. 1997. "A Chaotic Model of Innovative Search: Some Answers, Many Questions," *Organization Science* (8:5), pp. 528-542.

Krippendorff, K. 2013. *Content Analysis*: *An Introduction to its Methodology*, Los Angeles, London, New Delhi, Singapore: Sage.

Lauritzen, G. D., and Karafyllia, M. 2019. "Perspective: Leveraging Open Innovation through Paradox," *Journal of Product Innovation Management* (36:1), pp. 107-121.

Laursen, K., and Salter, A. 2006. "Open for Innovation: The Role of Openness in explaining Innovation Performance among U.K. Manufacturing Firms," *Strategic Management Journal* (27:2), pp. 131-150.

Laursen, K., and Salter, A. J. 2014. "The Paradox of Openness: Appropriability, External search and Collaboration," *Research Policy* (43:5), pp. 867-878.

Leiponen, A., and Helfat, C. E. 2010. "Innovation Objectives, Knowledge Sources, and the Benefits of Breadth," *Strategic Management Journal* (31:2), pp. 224-236.

Levinthal, D. A., and March, J. G. 1993. "The Myopia of Learning," *Strategic Management Journal* (14:S2), pp. 95-112.

Li, D., Eden, L., Hitt, M. A., Ireland, R. D., and Garrett, R. P. 2012. "Governance in Multilateral R&D Alliances," *Organization Science* (23:4), pp. 1191-1210.

Lind, J. T., and Mehlum, H. 2010. "With or Without U? The Appropriate Test for a U-Shaped Relationship*," *Oxford Bulletin of Economics and Statistics* (72:1), pp. 109-118.

Lopez-Vega, H., Tell, F., and Vanhaverbeke, W. 2016. "Where and How to Search? Search Paths in Open Innovation," *Research Policy* (45:1), pp. 125-136.

Lowry, M., Michaely, R., and Volkova, E. 2016. "Information Revelation Through Regulatory Process: Interactions between the SEC and Companies Ahead of the IPO," *SSRN Electronic Journal*.

Moss, T. W., Renko, M., Block, E., and Meyskens, M. 2018. "Funding the Story of Hybrid Ventures: Crowdfunder Lending Preferences and Linguistic Hybridity," *Journal of Business Venturing* (33:5), pp. 643-659.

Nadkarni, S., and Chen, J. 2014. "Bridging Yesterday, Today, and Tomorrow: CEO Temporal Focus, Environmental Dynamism, and Rate of New Product Introduction," *Academy of Management Journal* (57:6), pp. 1810-1833.

Nason, R. S., and Patel, P. C. 2016. "Is Cash King? Market Performance and Cash during a Recession," *Journal of Business Research* (69:10), pp. 4242-4248.

Nelson, R. R., and Nelson, K. 2002. "Technology, Institutions, and Innovation Systems," *Research Policy* (31:2), pp. 265-272.

O'Brien, J. P., and Folta, T. B. 2009. "A Transaction Cost Perspective on Why, How, and When Cash Impacts Firm Performance," *Managerial and Decision Economics* (30:7), pp. 465-479.

Peters, R. H., and Taylor, L. A. 2017. "Intangible Capital and the Investment-q Relation," *Journal of Financial Economics* (123:2), pp. 251-272.

Randhawa, K., Wilden, R., and Hohberger, J. 2016. "A Bibliometric Review of Open Innovation: Setting a Research Agenda," *Journal of Product Innovation Management* (33:6), pp. 750-772.

Ritala, P., and Hurmelinna-Laukkanen, P. 2013. "Incremental and Radical Innovation in Coopetition-The Role of Absorptive Capacity and Appropriability," *Journal of Product Innovation Management* (30:1), pp. 154-169.

Salge, T. O., Bohné, T., Farchi, T., and Piening, E. P. K. 2012. "Harnessing the Value of Open: The Moderating Role of Open Innovation Management," *International Journal of Innovation Management* (16:3), pp. 1-26.

Sirmon, D. G., Hitt, M. A., Ireland, R. D., and Gilbert, B. A. 2011. "Resource Orchestration to Create Competitive Advantage," *Journal of Management* (37:5), pp. 1390-1412.

Stanko, M., Fisher, G., and Bogers, M. 2017. "Under the Wide Umbrella of Open Innovation," *Journal of Product Innovation Management* (34:4), pp. 543-558.

Tan, D., and Mahoney, J. T. 2006. "Why a Multinational Firm Chooses Expatriates: Integrating Resource-Based, Agency and Transaction Costs Perspectives*," *Journal of Management Studies* (43:3), pp. 457-484.

Tarafdar, M., Gordon, S. 2007. "Understanding the Influence of Information Systems Competencies on Process Innovation: A Resource-based View," *The Journal of Strategic Information Systems* (16:4), pp. 353-392.

Teece, D. J. 1980. "Economies of Scope and the Scope of the Enterprise," *Journal of Economic Behavior & Organization* (1:3), pp. 223-247.

Teece, D. J., Pisano, G., and Shuen, A. 1997. "Dynamic Capabilities and Strategic Management," *Strategic Management Journal* (18:7), pp. 509-533.

Terwiesch, C., and Xu, Y. 2008. "Innovation Contests, Open Innovation, and Multiagent Problem Solving," *Management Science* (54:9), pp. 1529-1543.

Tong, T. W., Reuer, J. J., and Peng, M. W. 2008. "International Joint Ventures and The Value of Growth Options," *Academy of Management Journal* (51:5), pp. 1014-1029.

Tsai, W., and Ghoshal, S. 1998. "Social Capital and Value Creation: The Role of Intrafirm Networks," *Academy of Management Journal* (41:4), pp. 464-476.

Uotila, J., Maula, M., Keil, T., and Zahra, S. A. 2009. "Exploration, Exploitation, and Financial Performance: Analysis of S&P 500 Corporations," *Strategic Management Journal* (30:2), pp. 221-231.

Veer, T., Lorenz, A., and Blind, K. 2016. "How Open is too Open? The Mitigating Role of Appropriation Mechanisms in R&D Cooperation Settings," *R&D Management* (46:S3), pp. 1113-1128.

West, J. 2014. "Open Innovation," in *Open Innovation through Strategic Alliances*, R. Culpan (ed.), Palgrave Macmillan.

West, J., and Bogers, M. 2014. "Leveraging External Sources of Innovation: A Review of Research on Open Innovation," *Journal of Product Innovation Management* (31:4), pp. 814-831.

West, J., Salter, A., Vanhaverbeke, W., and Chesbrough, H. 2014. "Open Innovation: The Next Decade," *Research Policy* (43:5), pp. 805-811.

Wooldridge, J. M. 2013. *Introductory Econometrics*: *A Modern Approach*, Mason: Cengage Learning.