



A one-hundred-year structural topic modeling analysis of the knowledge structure of international management research

Mohamed M. Mostafa¹

Accepted: 26 September 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

International Management is a vast and multidisciplinary research domain that is heavily influenced by several other disciplines, such as Economics, Organizational Theory and Strategic Management. Based on 28,973 research articles, this study aims to analyze the knowledge structure of the international management domain from 1920 to 2019. Using computational text-based topic modeling analysis, we trace the evolution of international management knowledge by examining the major academic topics/latent themes discussed in the field. The study also diachronically visualizes the variations in topic prevalence over time. Our methodology is akin to “inductive mapping” as it is neither biased by our position nor it is guided by assumptions related to the topics we expect to find. Results indicate the existence of a wide variety of important research foci in the domain of international management. These include, among others, strategic alliances formation, international entry modes, corporate social responsibility, cross-cultural consumer behavior, technological innovation and entrepreneurship. Results also show that some topics such as “financial risk and return on investment” and “corporate social responsibility” show a declining time trend, indicating that academic research focusing on such topics was more likely to be published early on and less so recently. On the other hand, other topics such as “Emerging (East) Asian nations” and “global mergers and acquisitions” show an increasing trend, indicating that more papers were published recently. Taken together, although our findings might reflect the breadth and depth of research in international management, they might also suggest that the bounds of this field are not well defined.

Keywords Topic modeling · STM · International management · Text mining · Machine learning

✉ Mohamed M. Mostafa
mostafa@usa.com

¹ Gulf University for Science and Technology, West Mishref, Kuwait

1 Introduction

Scholars have regularly attempted to document the domain and the intellectual structure of international management research (Martinez and Toyne 2000). However, such attempts tended to either focus narrowly on one particular aspect of the field (Fan et al. 2021) or concentrate only on a short time span (Acedo and Casillas 2005; Pisani 2009; Werner 2002). With few exceptions, this type of research has traditionally employed content analysis in a systematic way to evaluate the content of international management research articles. However, the increased computational capacities coupled with the increasing availability of massive electronic archives have recently ignited the interest in computational text analysis techniques (Hagen Forthcoming). For example, a recent study using bibliometric analysis has reported 1171 published research articles dealing with topic modeling between 2000 and 2017 (Li and Lei 2021). In fact, the current moment has been termed “the computational turn” (Latzko-Toth et al. 2017, p. 199).

Topic modeling is a non-supervised machine learning technique that can detect hidden topics in a corpus without the need for an intervention from the researcher. The basic idea behind topic modeling is that each document is regarded as a mixture of topics, whereas each word has a certain probability of belonging to a specific topic (Hajdinjak et al. 2020). By “letting the text talk”, quantitative results obtained provide “an empirical basis for what might otherwise be informal claims about the discipline” (Malaterre et al. Forthcoming, p. 3). Thus, topic modeling results are, in general, free from bias stemming from the adherence to specific ideological perspectives (Jelveh et al. 2015). Although Baumer et al. (2017) found that topic modeling techniques can extract topics that are highly similar to that extracted by human experts, the former can produce reproducible results because of the exclusion of human coders (Symitsi et al. 2021).

Although topic modeling techniques have been applied to investigate a plethora of managerial topics such as corporate behavior (Bellstam et al. 2016), management science (Gatti et al. 2015), financial market efficiency (Huang et al. 2016) and monetary policy-making (Hansen et al. 2018), to the best of our knowledge, there has been no studies investigating the intellectual structure of international management research using topic modeling. In this research, we fill this research gap by examining the knowledge structure of international management research over one hundred years. In so doing, we believe that we have at least three major contributions to existing literature. First, in conducting our topic modeling analysis, we do not resort to a preconceived taxonomy of international management research as did Werner (2002) and Pisani (2009), who classified research in international management into twelve categories. Instead, we use an inductively grounded analytical approach to categorize international management research. Methodologically, we also contribute to existing research methodologies used in international management research by demonstrating how machine learning-based topic modeling can be used to detect the knowledge structure of the field. Thus, we contribute to the growing stream of research employing text-as-data methods (Blaydes et al. 2018). Finally, by analyzing the dynamics of topics over time, we identify which topics have endured or faded within the field of international management. More specifically, we aim to answer the following three major research questions:

RQ1. What are the major academic topics discussed in international management research during the last hundred years?

RQ2. What is the prevalence/distribution of the topics across the different subfields of international management?

RQ3. How do these research topics evolve across time?

The remainder of this paper is organized as follows. The next section deals with the literature review related to topic modeling. The following two sections describe the methodology and the results. The final section deals with the research implications and limitations. This section also explores future research avenues.

2 Related work

Topic modeling techniques have been applied to examine both structured and unstructured texts. More importantly, such techniques have been employed to investigate a broad array of social policy research questions such as public sentiments towards advocacy organizations (Bail et al. 2017), attitudes toward palliative care for dementia (McInerney et al. 2018) and gauging sentiments towards policy impacts (Flores 2017).

2.1 Structured texts

Although applying topic-based models to analyze massive textual corpora “is still in its infancy” (Gutmann et al. 2018, p. 284), topic modeling has been used in the literature mainly by digital humanity scholars to make sense of changing paradigms in the research field. This is usually achieved through the analysis of large-scale textual corpora such as examining all the papers published in *Science* over one hundred years (Blei and Lafferty 2007) or analyzing the abstracts of dissertations representing 240 universities over thirty years (McFarland et al. 2013). For example, in a study evaluating the evolution of economics from 1845 to 2013, Ambrosino et al. (2018) used topic modeling techniques to reveal the hidden thematic structure in economic research abstracted in the JSTOR database. The authors found that the field has grown exponentially, with 8220 articles published only between 2011 and 2013. By applying Latent Dirichlet Analysis (LDA) to track the macro-developments in the economics discipline, the authors constructed a map of the field over around two centuries. By so doing, it was possible to unearth the latent shifting structure of economics “in a time of (possible) fragmentation” (p. 329).

Kosnik (2015) used topic modeling techniques to chart the history and trends in economics based on 20,321 articles published in top-notch economics journals between 1960 and 2010. The author identified a short list of keywords representing well-delineated economic sub-fields. Results also showed that, apart from the decreased prevalence of macroeconomics, attention given to various economics sub-fields has remained relatively constant over time. Jelveh et al. (2015) analyzed a large corpus of economists’ scholarly writings from 1973 to 2011 to predict the individual political behavior of economists. Results revealed that the topics and results of the research are both influenced by the political ideology of the economists. Wehrheim (2019) applied topic modeling to analyze 2675 articles published in the *Journal of Economic History* between 1941 and 2016. Results showed that a “cliometric revolution” occurred in the 1960s based on the increasing dependence by economists on the use of quantitative/econometric methods to test economic theories.

Choi et al. (2017) used LDA to analyze the major research trends in the field of personal information privacy based on 2356 journal articles indexed in the Web of Science (WoS) between 1972 and 2015. Results showed that trending topics in the field are mostly related to social networks, new algorithms, Facebook and technology. Grothe-Hammer and

Kohl (2020) employed structural topic modeling (STM) to analyze the dynamic evolution of organizational sociology as a sub-field of sociology. In a similar vein, Jeyaraj and Zadeh (2020) used topic modeling and latent semantic indexing to chart the evolution of information systems (IS) research based on 2962 articles published in six leading IS journals between 2003 and 2017. The authors found that while topics such as information technology (IT) adaptation and IT development had endured over time, topics such as e-commerce and IT outsourcing had waned. Moro et al. (2019) used a topic modeling approach to trace and map the landscape of ethnic marketing research based on 239 articles indexed in the WoS and published between 2005 and 2015. Results revealed that the field is characterized by high conceptual heterogeneity, which might indicate “the lack of a clear, widely accepted definition” (p. 281). Bohr and Dunlap (2017) analyzed the major trends in the field of environmental sociology over twenty-five years (1990–2014) using STM. Results showed three major clusters: environment and society, social theory and natural resources.

Chen et al. (2020) combined STM with bibliometric network analysis to investigate the major trends in artificial intelligence-assisted human brain research over the last decade. The dataset used for analysis included 6317 papers published in the field of AI-assisted human brain research. The authors found that topic modeling approaches can help in determining the prominent research topics in this interdisciplinary field and chart how these research topics evolved over time. Similarly, Sharma et al. (2021) combined STM with bibliometric analysis to chart the landscape of information management research over fifty years. Based on 19,915 documents, the authors highlighted the most common themes in the field and revealed potential academic hotspots. Recently, Lemay et al. (2021) employed STM to analyze the evolution of both learning analytics and educational data mining based on a corpus of 681 abstracts representing the two fields. Results showed that the five extracted topics indicate little disciplinary differences between the two scientific fields in terms of research focus.

Sun and Yin (2017) used an LDA model to map the field of transportation research based on 17,163 articles published in twenty-two transportation journals between 1990 and 2015. Results revealed that the discovered topics match well-established subfields in transportation research. Topic distributions were also used to cluster journals and nations based on affinity/similarity. Analyzing topic dynamics over time showed that topics related to non-motorized mobility, travel behavior and sustainability have gained increased popularity over time. Similarly, Daenekindt and Huisman (2020) used STM to analyze the abstracts of 16,928 articles dealing with higher education from 1991 to 2018. The authors were able to trace how topics in the field have evolved over time. The authors also analyzed co-occurring topics, which resulted in revealing gaps in the literature, corroborating the claim that the field of higher education consists of an archipelago of “isolated islands.” Cirillo et al. (2020, p. 15) applied probabilistic topic modeling to classify 3364 multidisciplinary employee mobility articles over four decades into twenty topics. Topics detected highlight “the structural divide across disciplines.”

Kushkowsky et al. (2020) applied LDA approach to delineate the disciplinary boundaries of corporate governance research between 1990 and 2015 based on 10,532 articles extracted from the WoS database. Results revealed that the major topics in governance research are corporate governance theory, audit committees and control of family firms. More recently, Gurcan et al. (2021) used a topic modeling-based approach to analyze 41,720 research articles mapping the field of human–computer interaction between 1957 and 2018. The LDA-based approach detected twenty-one major topics, including online social communication, machine control systems and user interface design. Similarly, Malaterre et al. (Forthcoming) used LDA to analyze 15,897 philosophy of science articles over

eight decades (1934–2017). Results showed 25 distinct research themes and 8 clusters. Temporal analysis of the field revealed a relative decrease in topics dealing with logic and philosophy of physics and an increase in topics discussing the philosophy of mind and epistemology.

Grajzl and Murrell (2019) used STM to analyze all text documents representing Francis Bacon's opus. The corpus included around one million words in total. The authors were able to uncover sixteen topics representing the most prominent ideas of this key figure whose original thought helped spur modern economic development. The discovered topics support the idea that the common-law culture can help understand the seventeenth century English culture and can also help to explain the context of the early English political and economic development. This work was later extended to include 432 text documents totaling 1,320,262 words representing both Bacon and Coke's opus. Using STM, the authors were able to reveal the distinctive approach to the pursuit of knowledge followed by each preeminent English scholar (Grajzl and Murrell 2021).

Based on German title books between 1454 and 1600, Dittmar and Seabold (2015) used topic modeling techniques to detect characteristically Protestant and Catholic language texts. The authors were able to analyze the interplay between religious content and media competition during the era of the Protestant Reformation in Germany. Similarly, Blaydes et al. (2018) used STM to analyze a medieval period corpus comprising Muslim and Christian political advice texts. The authors also examined how the emphasis on advice has changed over time. For example, an inflection point in Muslim texts was identified between the eleventh and the thirteenth centuries. This inflection point was linked to the Turkic and Mongol conquest of the Islamic world. A decline in the prevalence of religious appeals from the Middle Ages to the Renaissance was detected in Christian political advice texts. Lucas et al. (2015) also used topic modeling techniques to analyze Islamic *fatwas* to determine whether the clerics issuing such religious edicts are Jihadists or not. Results revealed crucial topical differences between the two groups. For example, Jihadists mostly focus on topics dealing with "Killing" and "fighting", whereas non-Jihadists tend to focus on topics such as "Ramadan" and "prayer."

Based on 351 articles published in the academic journal *Accounting History* over twenty years from 1996 to 2015, Ferri et al. (2018) used STM to analyze the relative prevalence of research areas and their evolution over time. Based on a corpus comprising 1,319,776 words, the technique used was able to uncover overlooked topical areas and refine existing categorizations. The authors also found a recent shift from the focus on the technical core of accounting to more fragmented approaches. Kang et al. (Forthcoming) used STM to identify 37 key thematic topics representing 1516 articles published in the *Nonprofit and Voluntary Sector Quarterly* between 1972 and 2019. The authors found that while there was a general trend of continuity, some topics showed a decline and fewer showed an increase over time.

Roose et al. (2018) also used topic modeling to analyze 6965 articles published in a leading art magazine between 1991 and 2015. The analysis revealed that the prevalence of major contemporary art criticism topics has barely changed over the study period. A notable exception relates to the "social turn", a concept that pays substantial attention to social forms and content. Similarly, Priva and Austerweil (2015) applied topic modeling methods to analyze titles and abstracts of thirty-four years of *Cognition* articles. Results showed that while articles dealing with moral cognition and eye-tracking had surged, topics related to developmental cognition remained stable over the study period. Cheng et al. (2018) combined topic modeling with bibliometric analysis and visualization techniques to analyze the ecology, environment and poverty nexus. Based on 4335 articles abstracted

in the WoS, the authors found that topics related to ecosystem, health risk, women and inequality represent the main concern in the interplay between poverty and the environment. Winson-Geideman and Evangelopoulos (2013) employed a topic modeling approach to analyze the knowledge structure of published real estate work based on four leading real estate academic journals over forty years. Utilizing the Singular Value Decomposition method within the Latent Semantic Analysis (LSA), the authors were able to extract socially-constructed components of meaning within the field of real estate.

Kusters and Garrido (2020) used STM to examine the image of Spain and the ‘European South’ in the German newspaper *Die Zeit* based on 2443 news articles published over more than sixty years between 1946 and 2009. The analysis allowed the authors to characterize a shift from tourism topics during the 1960s (the Golden Age of Germany’s reconstruction following WW II) to issues related to economic competitiveness and European integration in the 1990s. Similarly, Ploessl et al. (Forthcoming) used topic modeling to analyze 168,012 articles published in a leading German real estate newspaper between 1999 and 2019. Topics revealed that articles related to urbanization and globalization represent the lion’s share of news coverage. Newman and Block (2006) used a probabilistic topic modeling approach to analyze a corpus of 80,000 articles and advertising texts published in the *Pennsylvania Gazette* between 1728 and 1800. By fitting a latent semantic model with forty topics, the authors found that most of the topics dealt with political and economic issues. Most prevalent ads dealt with indentured servants and escaped slaves. Time shifts detected indicated an exponential increase in governmental issues between the 1760s and the 1790s. For example, discussions of clothes were tracked to the ebbs and flows of the US imports during this period.

Franzosi et al. (2012) used a topic modeling approach to analyze the US newspapers’ accounts related to lynchings between 1875 and 1930. Results revealed a correlated direct network representing the spatial distribution and the chronology of lynching in the US. Using LDA techniques, Mueller and Rauh (2019) analyzed the changing themes in newspaper articles published in the English-speaking world between 1975 and 2015 to predict the timing of armed conflicts. Terman (2017) investigated journalistic texts appearing in the *Washington Post* and the *New York Times* between 1980 and 2014 to scrutinize portrayals of Muslim women in the U.S. media. Results revealed a systematic bias towards Muslim women as stories about them tended to focus on the violation of their rights, while stories about non-Muslim women tended to focus on several other topics. Recently, Park et al. (2021) employed STM to identify major topics related to tourism demand based on CNN and *China Daily* articles. Results based on topics extracted highlighted diverse issues related to Mainland Chinese tourist arrivals, including foreign relations, international events and social issues.

Using the 2005 version of the *Congressional Record*, Gentzkow and Shapiro (2010) used a topic modeling approach to identify phrases most likely to be used by a particular political party. The authors also compared the obtained phrases against a corpus representing the major US newspapers to detect the newspaper’s “slant” or ideological orientation between left and right. In a similar vein, Jensen et al. (2012) used a corpus of Congressional debate to examine more than a century of political partisan polarization. Bagozzi and Berliner (2018) applied STM to analyze the US State Department’s Annual Country Reports on Human Rights Practices. The corpus included 6298 texts ranging between 1977 and 2012. Results revealed a plausible set of topics including, among others, human trafficking, prison conditions, labor rights, freedoms of expression and forced disappearances. Results also revealed a steep decline in some topics such as the topic labeled “economic systems” from 1977 until roughly 1990.

Hajdinjak et al. (2020) used STM to investigate migration policy framing at the US House of Representatives and the Canadian House of Commons between 1994 and 2016. Results revealed considerable differences in the framing emphasis in migration-related topics between the two ideological blocks. For example, while Democrats in the US and liberals in Canada emphasize welfare and humanitarian aid, conservative blocks emphasize the security and legal aspects of migration. Altaweel et al. (2019) used a topic modeling approach to examine government documents including information dealing with ecological disturbance. Results revealed that documents dating back to the 1960s had focused on topics such as outbreak factors and tree population, whereas recent topics focus on terms such as “fire” and “mortality.” Bao and Datta (2014) employed a topic modeling approach to detect financial risk disclosures from companies’ annual reports. Recently, Sun and Wang (Forthcoming, p. 1) used STM to analyze webpage narratives related to 200 dual languages (DL) programs in the US. Topics extracted show that DL programs are promoted “as a model that brings cognitive and employment benefits” to students. Similarly, Goyal and Howlett (2021) used topic modeling to analyze the content of 13,000 Covid-19 policies worldwide. Results show significant cross-country variation in terms of policy intensity, density and balance of policy mixes.

Applying both STM and LDA topic modeling approaches, Maerz and Schneider (2020) examined 4740 speeches by forty democratic and autocratic political leaders between 1999 and 2019. Results revealed that liberal and illiberal leaders around the globe present themselves differently both to their national and international audiences. For example, the authors found that while the topic of “collective memory” in democracies is dealt with from the perspective of the holocaust, death and freedom, the same topic is closely linked to national identity and authoritarian ideologies in the case of autocratic leaders’ speeches. Dybowski and Adammer (2018) also combined correlated topic modeling and sentiment analysis to investigate how the US presidential speeches related to tax communication affect the economy. Results showed that optimistic tax policy statements in the presidential speeches stimulate both consumption and investment. Econometric results also revealed that consumers’ sentiments react positively to optimistic tax news. Quinn et al. (2010, p. 224) used a total of 118,065 legislative speeches to infer the “relative amount of legislative attention paid to specific topics.” Stewart and Zhukov (2009) used topic modeling techniques to analyze around 8000 speeches by Russian political and military elites between 1998 and 2008. Results showed that topic modeling can be used as a useful tool in understanding the public debate regarding the use of force as a foreign policy instrument. Recently, Ovadek (2021) used topic modeling to classify the content of 628 presidential speeches in Slovakia from 1993 to 2020. The author tracked topical variation over time for different presidents and was able to detect salient issues such as intra-executive conflict and agenda shifts.

2.2 Unstructured texts

Prior research using topic modeling to analyze unstructured texts has focused on four major research streams, namely consumer-generated online reviews/feedback (Ding et al. 2020; Hu et al. 2019; Lang et al. 2020), open-ended questionnaires (Lee and Kolodge 2020; Chung et al. Forthcoming), tweets and social media posts (Kwon et al. 2019) and online fora comments (Tornberg and Tornberg 2016). For example, Sanchez-Franco et al. (2021) used both STM and LDA to analyze 31,026 consumer-generated comments on Amazon’s Echo and Google Home products spanning the period between 2017 and 2019. Results

showed that consumer comments are largely driven by hedonic and utilitarian gratification, social norms and facilitating conditions. Guo et al. (2017) applied topic modeling to analyze 266,544 consumer-generated online reviews representing 25,670 hotels worldwide. LDA detected nineteen salient dimensions that can be utilized by hotels to manage their interaction with visitors.

Luo et al. (2020) utilized topic modeling to analyze 43,869 online tourist reviews related to Disneyland theme parks in Hong Kong, Paris and California. The most important topics were associated with happy experiences, shopping options, booking systems, value for money and the food court. Using a topic modeling approach, Heng et al. (2018) analyzed 22,424 customer reviews related to coffee. Results show four major topics, namely service, physical features, flavor features and subjective expression. Srinivas and Rajendran (2019) scrapped thousands of publicly available students' reviews from online university sources. The authors used topic modeling to classify reviews into themes such as academic support, financial aid, job prospects and social life. Fagernas et al. (2021) analyzed 1379 online consumer reviews related to virtual reality (VR) relaxation applications. A topic modeling approach was used to detect salient themes related to VR, including gamification elements and immersion and/or related sensory topics.

Korfatis et al. (2019) used both STM and LDA approaches to analyze 557,208 airline passengers' online reviews. Major topics detected include value for money, baggage policy, delays and refund policy. Liu et al. (2021) used STM to analyze responses to 1452 open-ended questions in a survey examining landscape preferences in Norway. Results revealed that consumer choices were driven mainly by themes such as landscape openness, concerns regarding deforestation, biodiversity issues and aesthetic properties. In a similar vein, Knudson et al. (Forthcoming) employed STM to analyze 1500 written responses in a survey examining Norwegian citizens' trust in news media. Results showed that responses may be summarized in four major topics: bias and independence, truthfulness, objectivity and professionalism.

Fresneda et al. (2021) showed that STM technique can be used as a market segmentation tool for the non-profit sector. The authors analyzed 8418 responses to an open-ended question in a nationally representative online sample to extract sixteen major topics related to the not-for-profit segment of the market. Major topics include community service, ethical consumption and welfare programs. Similarly, Bennett et al. (Forthcoming) employed STM to analyze the open-ended answers of 544 respondents regarding their attitudes towards food banks. Major topics detected relate to "deservingness", "vulnerability" and "unfair society." Savin et al. (2021) used topic modeling to examine 1822 open-ended textual responses concerning the link between "economic growth" and "green growth." Major topics detected include social and environmental problems, material wealth, green technologies and sustainability.

Han et al. (2021) used web scraper software to collect 129,965 Covid-19 pandemic tweets posted by UK citizens and news media outlets in April 2020. STM detected wide diversity in topics dealt with by both groups, indicating different agenda settings during the pandemic. For example, while tweets posted by news media focused on topics such as death statistics and international Covid-19 news, citizens focused on themes such as social distance and quarantine/lockdown measures. Similarly, Fino et al. (Forthcoming) used topic modeling to explore gambling addiction topics and sentiment during the Covid-19 pandemic based on tweets posted in April 2020. Topics detected were related to gambling addiction perception, risks associated with gambling, and forms of gambling during the pandemic. Rodriguez and Storer (2020) also used STM to analyze tweet posts to explore why people decide to leave or stay in abusive relationships.

Melton et al. (Forthcoming, p. 1) also used topic modeling techniques to analyze posts on Reddit, a social media platform. More specifically, the authors collected textual data posted by thirteen Reddit communities focusing on Covid-19 between December 1, 2020, and May 15, 2021. Results show that community members have “focused on side effects rather than outlandish conspiracy theories.” Puschmann et al. (2020) compared 315,522 comments posted on the Facebook page of two German right-wing movements. Results revealed salient topics such as migration, opposition to Islam and government policies. Barry et al. (2018) used LDA to examine all tweets posted by thirteen alcohol brands between 2010 and 2017. Results showed that each brand uses a unique approach to marketing based on the brand personality itself. Tweets also revealed that only a few posts were allocated to encourage moderate drinking. Heiberger et al. (Forthcoming) used STM to examine how political actors and news media contribute to political fragmentation based on around half a million election-related tweets extracted in 2017. Results revealed that political actors contribute more to fragmentation as they focus on contentious topics while articulating their policies for different constituencies.

Schmiedel et al. (2018) used topic modeling to examine 428,492 Fortune 500 reviews posted by employees on the online platform Glassdoor. The authors were able to identify the most important topics that mattered to employees. Results also showed that there is an association between such topics and employees’ perceptions of organizational culture. Similarly, Sainju et al. (2021) used STM to explore 682,176 employees’ reviews of Fortune 50 posted on the Indeed.com platform. Results showed that salient topics related to employees’ satisfaction include monetary benefits, top management support and work environment. Symitsi et al. (2021) used STM to analyze 349,550 employee reviews posted on the Glassdoor platform. Results indicated that unstructured data have informational value for firms’ internal and external stakeholders. Stamolampros et al. (2019) used topic modeling to examine 297,933 employee comments posted on an online forum to investigate factors affecting employees’ satisfaction in the US tourism and hospitality sector. The authors found that the major topics influencing satisfaction and turnover include leadership and cultural values. In a similar vein, Jung and Suh (2019) analyzed employees’ comments on a major Korean platform to examine major themes related to satisfaction. Using LDA method, the authors found that employees’ satisfaction is driven mainly by senior management support and compensation.

From this extensive literature review, we see that there is no prior work done to apply topic modeling to track the evolution of the field of international management over a long period of time. In this research, we make the first attempt to apply “text-as-data” techniques to examine the knowledge structure and the evolution of international management over one hundred years. We also hope that this work could stimulate more discussion about this multidisciplinary field of research.

3 Method

3.1 Database and search terms

Following Ambrosino et al. (2018), published documents on international management were retrieved from the digital library JSTOR. The JSOR provides, upon agreement, its Data for Research (DfR), which includes the Optical Character Recognition (OCR) full text of most articles along with metadata. This contrasts sharply with other databases

providing only article abstracts such as the WoS, Scopus or PubMed. Thus, DfR can be utilized for applying topic modeling techniques to the full content of international management articles. Only research article titles were searched using all possible variations of the terms “International” AND “Management” OR “Business.” This decision was taken to avoid false-positive results since a title-specific search was found to minimize loss of sensitivity and increase recovery rate (Alexandre et al. 2015). The search process was limited to documents written in English and we chose 1920 as the date of reference because we intended to evaluate the knowledge structure and the evolution of the field of international management over one hundred years (January 1, 1920, till December 31, 2019). Using the search terms indicated yield 28,973 research articles. We excluded 19,810 documents representing book reviews, article commentary, review articles, editorials, etc. The distribution of international management research articles within the JSTOR corpus is plotted in Fig. 1. As can be seen, there has been an exponential growth in articles dealing with international management, with a particularly strong rise towards the end of the last century and early 2000s.

3.2 Data pre-processing

Following the recommendations of Denny and Spirling (2018), the obtained documents were subject to pre-processing and/or cleaning following the standard procedures associated with bag-of-words (Ovadek 2021). The obtained texts were also rearranged to make the analysis more statistically tractable. Punctuation, capitalization, digits and whitespaces were removed. Terms were tokenized and lemmatized and a stemming procedure was followed so that words with the same root were grouped together. For instance, the words “investment”, “investor” and “investing” were all classified under

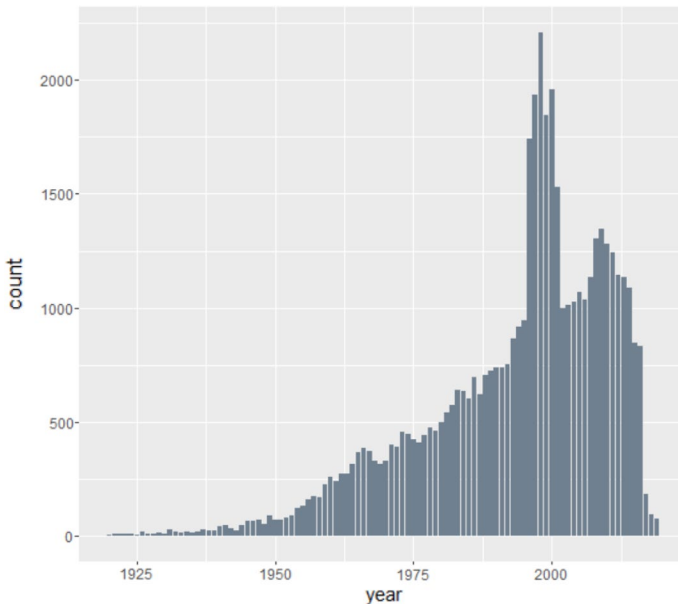


Fig. 1 Distribution of international management research in the JSTOR database

“invest-” By removing the ends of words, stemming aims at reducing complexity without sacrificing important information through the reduction of the total number of unique words. Stop words such as “and”, “or” and “is” were also eliminated since such words have high frequency but no substantive meaning. Additionally, words with low frequency, which appeared less than three times in the entire corpus were removed. This decision was made to make sure that extremely rare vocabulary (often words with typos) does not confound the findings. Several packages within the R version 4.1 environment (R Development Core Team 2021) were used to conduct the analysis, including *stm*, *tidystm*, *tidyverse*, *tm*, *wordcloud*, *qgraph*, *huge*, *LDA*, *LDavis* and *quenteda*.

3.3 STM

First developed by Roberts et al. (2013, 2014), STM is an unsupervised machine learning algorithm that attempts to detect the “hidden structure” in textual data. Unsupervised algorithms learn and discover the hidden or underlying structure in a given corpus and produce the top terms representing each topic or topics (Greene et al. 2014). Unlike LDA and similar supervised topic modeling algorithms, STM “allows correlated topics to be identified” (Fresneda et al. 2021, p. 796). This is a critical ability because social science topics tend to commonly correlate with each other. Blei and Lafferty (2007, p. 18) noted that the failure of supervised algorithms to deal with correlated topics might lead to an “unrealistic modeling assumption that the presence of one topic is not correlated with the presence of another.” In fact, the major assumption of STM is that each document is composed of multiple topics in different proportions. This implies that the estimated topics are allowed to vary based on designated criteria. Furthermore, STM allows for the inclusion of metadata/covariates in the model estimation and/or inference process.

One common challenge in any STM application is the selection of the appropriate number of topics (k). Defining a very small number of topics might result in inferring broad and rather unspecific categories. On the other hand, selecting a very large number of topics tend to create multiple marginal clusters of insignificant topics. Savin et al. (2021) noted that the larger the k , the harder it is to interpret the model since a large k value complicates the model. Roberts et al. (2016) argued that the optimal number of topics should be determined using the highest held-out likelihood based on a combination of semantic coherence and exclusivity. Known in the literature as FREX, this criterion is based on the weighted harmonic mean of a word’s rank in terms of Frequency (FR) and Exclusivity (EX) in a k -topic solution. The held-out likelihood concept is based on the idea that the information learned from a certain corpus should work well when applied to an unseen document. Thus, a better-performing model would have a higher held-out probability (Lemay et al. 2021). Taddy (2012) argues that residuals can also be used as a diagnostic tool in STM as an overdispersion of residuals might indicate that the “true” number of topics is higher than the number suggested by the model. Semantic coherence can be used to evaluate the quality of topics estimated by the model (Pandur et al. 2020). Developed by Minmo et al. (2011), semantic coherence is akin to the pointwise mutual information criterion (Lau et al. 2014) and is maximized when the “most probable words in a given topic frequently co-occur together” (Sainju et al. Forthcoming). Finally, the lower bound concept can be used to trace the convergence of the model. This is achieved when a small enough change between iterations occurs (Roberts et al. 2014).

4 Results

4.1 Number of topics k

The first step in initializing any STM application is to determine the “optimal” number of topics. It should be emphasized that there is no “right” number of topics as this depends on both the analytic utility and the interpretability with regard to the research question (Grimmer and Stewart 2013). Following Daenekindt and Huisman (2020), we determine the number of topics using the spectral method of initialization. Compared to other methods, this method is faster and more accurate and it guarantees to obtain global optimal model parameters (Roberts et al. 2016). Similar to the approach followed by Tvinnereim et al. (2017), we run the STM by specifying the number of topics anywhere from 5 to 40, in 5-topic increments. Relative goodness of fit results for each number of topics is shown in Fig. 2.

Figure 2 shows that between 15 and 20 topics produce relatively high held-out likelihood, high semantic coherence, maximized lower bound and low residuals. Having completed the first stage and established the “optimal” latent topics range generated from the 28,973 research articles, we relied on Griffiths and Steyvers’ (2004) two-stage algorithm to determine the “optimal” number of topics. By conducting model-fit statistics/log-likelihood for each topic in the range from 15 and 20, the log-likelihood was maximized at

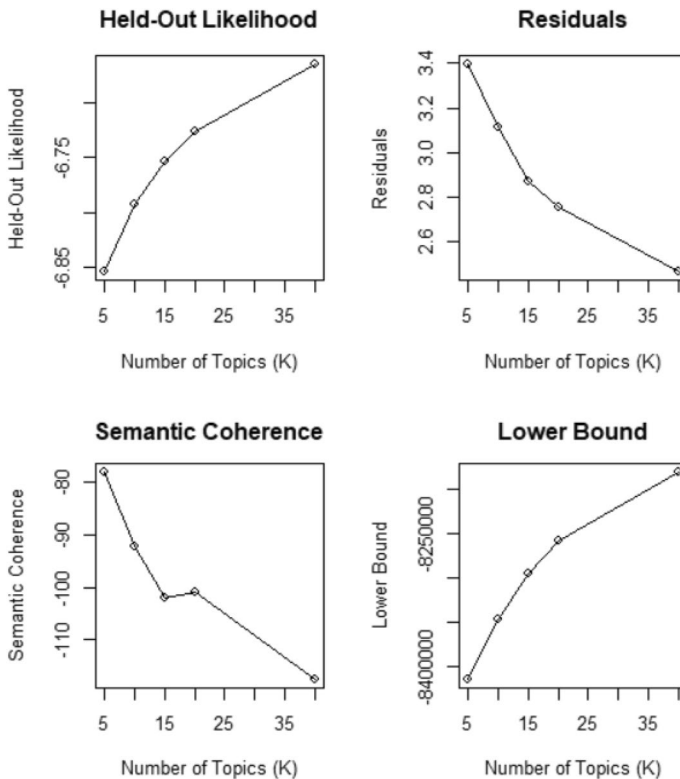


Fig. 2 Determining the optimal range of latent topics through model-fit statistics

$k=18$. Guided by this algorithm, we conclude $k=18$ would be the “optimal” number of latent topics. Kusters and Garrido (2020) warned against obtaining the “optimal” number of topics only in a statistical sense by simply running the algorithm several times and comparing the respective results. Thus, to decide on the most coherent and interpretable output, we also calculate the semantic coherence and exclusivity for different topics. Figure 3 shows that both coherence and exclusivity are maximized between $k=15$ and $k=20$, confirming the robustness of our statistical results.

This result implies that the model not only fits the data but also comprises semantically interpretable topics with keywords that are unlikely to overlap with keywords from other topics.

4.2 Topic prevalence

Topic prevalence refers to “the relative importance of the various topics in the corpus” (Grajzl and Murrell 2019, p. 123). This importance is measured by the proportion of a document devoted to a certain topic. Figure 4 shows the expected values for topic proportions in the international management corpus. The size of the bar adjacent to each topic represents the probability of selecting a random word generated by that particular topic from the whole corpus. Not surprisingly, topic five deals with culture and (cross)-cultural consumer behavior is the most prevalent in the corpus with about 10% of the international management corpus-wide occurrences.

Lu (2003, p. 196) noted that this research stream has been very rich as it “involves both the identification of dimensions of culture relevant to international operations, as well as tests of how cultural variance influences firm strategy and performance.” In fact,

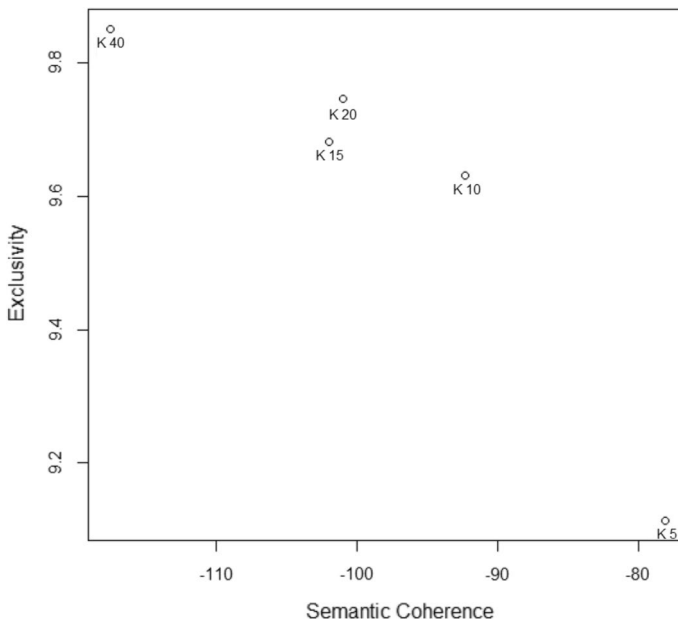


Fig. 3 Determining the trade-off between semantic coherence and exclusivity

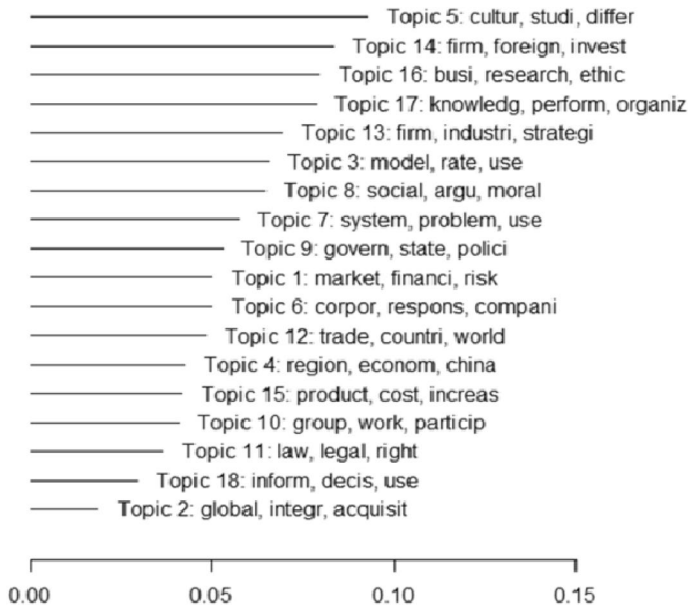


Fig. 4 Expected topic proportions and top three high probability words associated with each topic

Hofstede's (1980, 1991) work in cultural differences among nations has come to represent a milestone in international management research. This work was followed by an intense debate among international management scholars regarding the difficulties and dangers of trying to blindly transfer managerial theories and practices abroad (Fey and Denison 2003).

Following Sharma et al. (2021), we used the maximum-a-posteriori (MAP) algorithm to estimate document-topic loadings. The resulting histogram is shown in Fig. 5. In this graph, the median estimate of document-topic proportions is shown in a dashed red line. From the graph, it is clear that each extracted latent topic is hardly related to several research documents, which is clearly consistent with the statistical mixture hypothesis, implying that "each document is a probabilistic mixture of a limited number of topics where the probability of key topics will be high and non-key topics will be low" (Sharma et al. 2021, p. 8). Thus, the graph further confirms the quality of the eighteen topics extracted from the international management research corpus.

To facilitate an aggregate analysis, it is common in the literature to assign a single label to each topic. Figure 6 shows an example of the top twenty FREX words defining the first four topics. Since these words are highly exclusive to each topic, they were used to develop the semantic labels for each topic as shown in Table 1.

Our results indicate the existence of a wide variety of important research foci in the domain of international management. These include, among others, strategic alliances formation, international entry modes (FDI, joint ventures, global mergers and acquisitions), corporate social responsibility, cross-cultural consumer behavior, technological innovation and entrepreneurship, expatriation and international team-building skills, international financial risk management, transaction cost and supplier efficiency, legal settlement of disputes, emerging markets, anti-corruption policies and moral and ethical issues. Although this finding might indicate the breadth and depth of research in international management, it might also suggest that the bounds of this field are not well defined (Ricks et al. 1990).

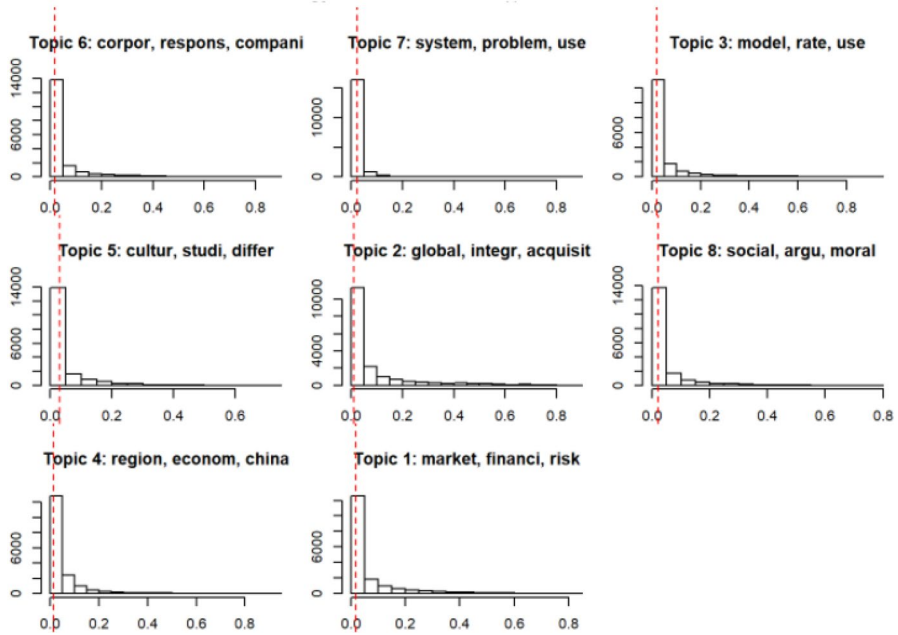


Fig. 5 Estimates for selected document-topic loadings (MAP method)

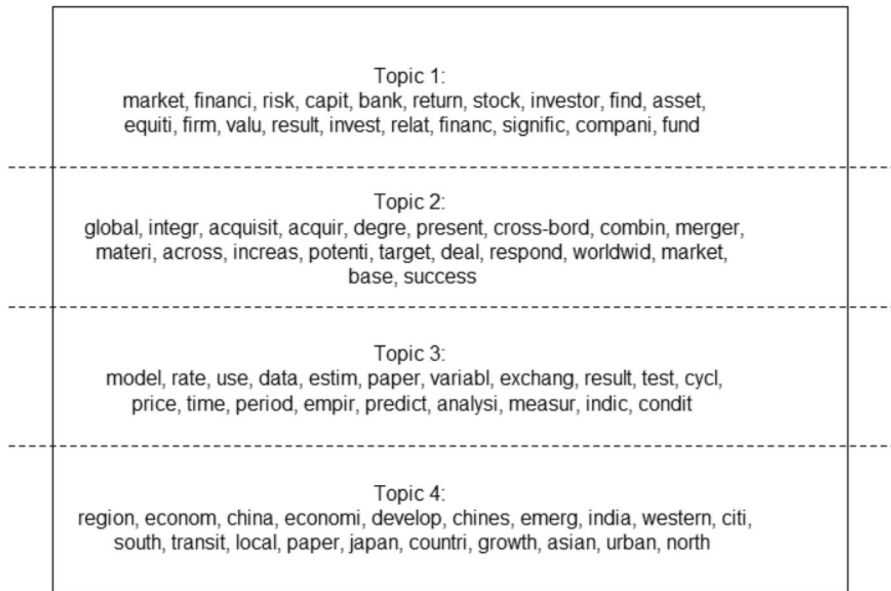


Fig. 6 Top twenty FREX words associated with the first four topics

Table 1 International management research topics and highest FREX stem words in each topic

Topic	Description	FREX
Topic 1	Financial risk and return on investment	Risk, stock, bank, financi, return, investor, asset
Topic 2	Global mergers and acquisitions	Integr, global, merger, acquir, acquisit, cross-bord, materi
Topic 3	Modeling cycles and estimating inflation rates	Cycl, estim, shock, rate, forecast, inflat, model
Topic 4	Emerging (East) Asian nations	Asian, asia, east, citi, korea, western, india
Topic 5	(Cross)-cultural consumer behavior	Cultur, brand, percept, behavior, perceiv, consum, attitud
Topic 6	Corporate social responsibility	Corpor, csr, stakehold, environment, respons, code, organis
Topic 7	Systems and mathematical modeling	Problem, comput, solut, plan, applic, system, solv
Topic 8	Moral and ethical issues	Moral, societi, discours, contemporari, immigr, democraci, religi
Topic 9	(Anti)-corruption policies	Govern, corrupt, polit, polici, british, state, offici
Topic 10	Expatriation, gender and team-building skills	Team, famili, women, job, expatri, skill, worker
Topic 11	Legal settlement of disputes	Law, legal, court, enforce, rule, disput, legisl
Topic 12	World trade and liberalism	Trade, union, european, world, labour, energi, liber
Topic 13	Technological innovation and entrepreneurship	Innov, technolog, competit, cluster, internat, entrepreneuri, smes
Topic 14	Subsidiaries, FDI and joint venture formation	Subsidiari, fdi, ventur, host, entri, export, foreign
Topic 15	Transaction cost and supplier efficiency	Product, tax, cost, incom, effici, supplier, suppli
Topic 16	International education	Student, journal, univers, school, academ, educ, teach
Topic 17	Strategic alliances and organizational learning	Allianc, knowledg, partner, organiz, trust, strateg, learn
Topic 18	Language and communication	Inform, formal, languag, decis, communic, english, make

model (CAPM) by Sharpe (1964) represents a major breakthrough in this regard as this model postulated that investors are risk-averse and would therefore seek a premium for assuming risk (Remmers 2004).

Figure 8 plots a comparative graph between topics six (corporate social responsibility) and fifteen (transaction cost and supplier efficiency). In this graph, the size of the stem word is proportional to its frequency, whereas the distance between any pair of stem words determines their closeness to a specific topic. Although each word's position on the vertical axis is random, the horizontal line determines whether a stem word falls either in the extreme left or right and thus pertains to one topic, or it occupies a central position and thus it appears as much in the two topics. For example, it is clear from the figure that the topic labeled “corporate social responsibility” emphasizes terms related to stakeholders and the environment, while the “production efficiency” topic frequently features words related to production and cost.

Corporate social responsibility has been investigated in the international management literature for decades (Bowen 1953; Carroll 1999). Egri and Ralston (2008) noted that corporate social responsibility comprises both social and environmental components, whereas the latter component increasingly gained momentum early in the 1990s (Hart 1995; Shrivastava 1996; Starik and Rands 1996). From the graph, we see that both components are fundamentally interwound as the topic dealing with corporate social responsibility (topic six) includes terms representing both components. On the other hand, the topic dealing with transaction cost and supplier efficiency (topic fifteen) finds its roots in the pioneering work of Coase (1937) who suggested that firms and markets are simply alternative methods of organizing production efficiently. The transaction cost theory gained little attention in the international management literature until the 1970s, when it was “rediscovered” by Williamson (1975, 1979, 1985).

4.3 Topic correlations

Unlike LDA, STM can explicitly incorporate the possibility of correlated topics across documents. This aspect is used here to explore the interconnectedness among the

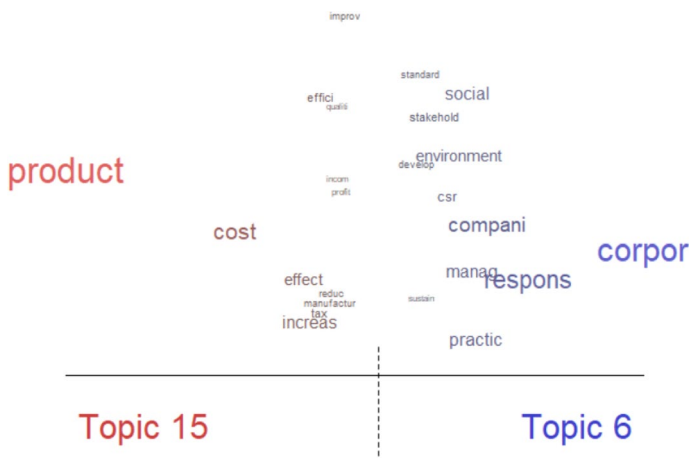


Fig. 8 The comparative graph between topics six and fifteen

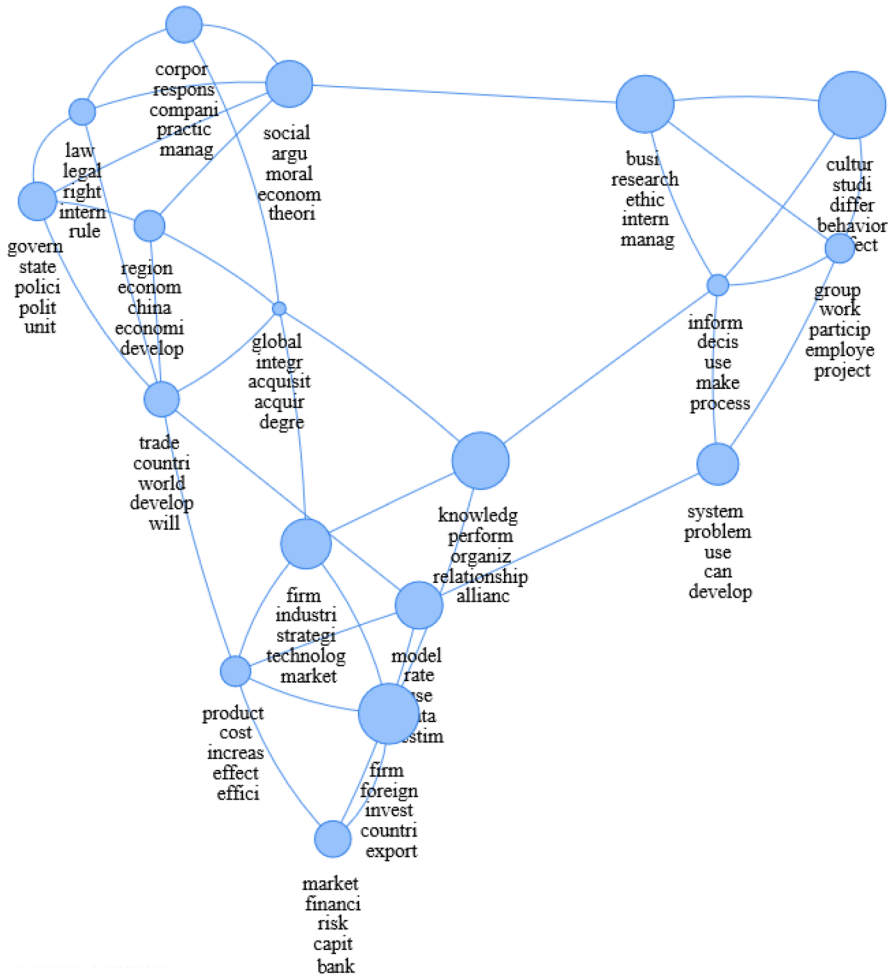


Fig. 9 Network of correlations between international management research topics

various topics in the international management corpus. Figure 9 depicts the network of topic correlations. In this graph, each topic is represented by a circle, whereas its size is proportional to the topic proportion in the entire corpus. Lines indicate correlations between topics, whereas a shorter link between topics indicates a stronger correlation between them. For example, we see close links between the topic discussing corporate social responsibility and the topic dealing with moral issues, while the same topic of corporate social responsibility is weakly related to global mergers and acquisitions. The correlations network was constructed using the Fruchterman and Reingold (1991) algorithm, which “causes groups of nodes that are more tightly connected to be closer together, with disconnected and low-degree nodes being placed on the outskirts of the plot” (Adhikari and Dabbs 2018, p. 237). It has been demonstrated that this algorithm, which uses attraction-repulsive forces to model nodes’ distances, is akin to the traditional multidimensional scaling (Borg and Groenen 1997).

From the graph we see that the “strategic alliances and organizational learning” topic is centrally located within the international management research network, indicating that this topic may be regarded as influential in the network. Park et al. (2015) argued that by forming bridges between other key clusters, centrally located topics may play an important role in information diffusion and brokerage through the network.

Interestingly, Werner (2002, p. 287) found that “the area of strategic alliance relationships are [*sic*] the most researched sub-topic” in international management research. This area includes strategic alliance formation and relationships, business networks and outcomes of strategic alliances. Classical research in this area includes alliance stability (Celly et al. 1999), continuity (Dyer and Chu 2000), negotiation tactics (Rao and Schmidt 1998), value creation (Holm et al. 1999), network responsiveness (Zaheer and Zaheer 1997) and changes in knowledge structure (Chikudate 1999). Research on international strategic alliances has benefited from the application of organizational learning theory (Makino and Delios 1996; Parkhe 1993; Yan and Gray 1994) and by conceptualizing the strategic alliance as a learning race (Hamel 1991; Inkpen and Beamish 1997). Many international management researchers have applied the organizational learning theory to explore the mechanisms through which learning occurs in international strategic alliances (Arino and de la Torre 1998; Inkpen and Crossan 1995; Simonin 1999).

Another way to map the structure of international management research is to reveal which topics tend to be combined together in research articles. This is akin to examining “how (dis)similar the topics are” (Ovadek 2021, p. 227), which can be done using clustering techniques. Figure 10 shows the dendrogram obtained by conducting a hierarchical clustering on the international management topics using the Euclidian distance within Ward’s minimum variance method. Since distance implies similarity between topics, the closer the

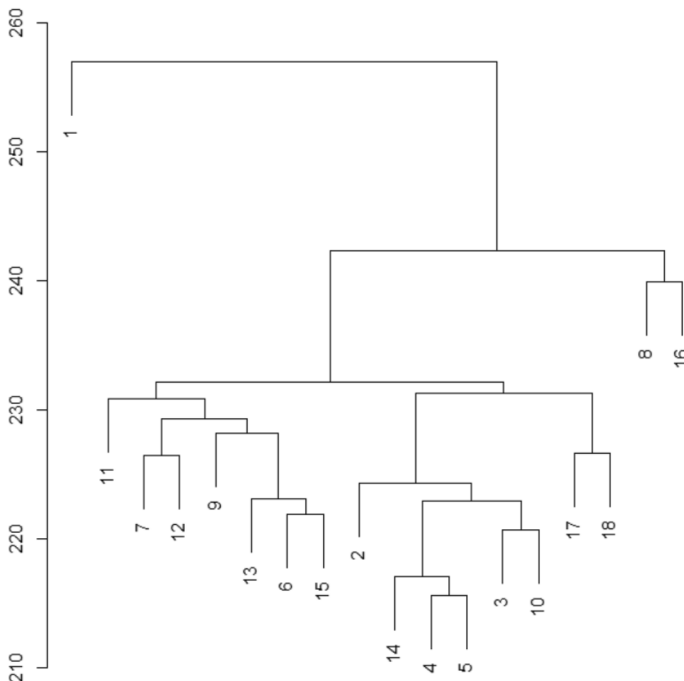


Fig. 10 Hierarchical clustering dendrogram of international management research topics (Ward’s method)

topics the more similar they are in terms of their distribution over the corpus. Thus, topics that cluster together might be regarded as topics that are most likely combined together in research articles. For instance, topic 12 “world trade and liberalism” and topic 7 “systems and mathematical modeling” are very close to each other. This is an interesting result as it corroborates Wehrheim (2019) thesis regarding the “cliometric revolution” which occurred in the 1960s and resulted in an exponential increase in the use of quantitative/econometric methods to test international trade and other economic theories. On the other hand, the topics “financial risk and return on investment” and “international education” are very far away from each other, indicating that both topics are rarely combined together in international management research.

The dendrogram is also a valuable tool in revealing the different “islands” in international management research. For example, the topic related to culture (topic 5) is typically addressed at the microlevel in international management research (Adler 1983; Child 1981), whereas the importance of political policies such as the anti-corruption policies (topic 9) is usually investigated at the macro level (Boddeyn and Brewer 1994). However, based on the distance between the two clusters in Fig. 10, it appears that the two topics are rarely investigated together, which implies that this might be a potential area of research in international management. Stening and Skubik (2007, p. 112) argued that “in a field such as international management research, ethics and politics are largely inseparable.” However, a closer look at Fig. 10 shows the lack of research combining both topics. This is remarkable given the fact that while the two topics are analytically distinguishable, “the politics and ethics of cross-cultural research are tightly interwoven” (Warwick 1980, p. 321). Thus, this appears to be a potential area of research in international management research. One would also expect that international environmental variables such as factors related to exchange rate risk and other economic factors (topic 1) affect the knowledge transfer and the management of expatriates (topic 10). However, the large distance between the two clusters in the graph indicates that the two topics are rarely combined together, which might indicate a potential research area within international management research.

4.4 Temporal dynamics

Having analyzed the international management topics’ content, we turn now to examine the way topics evolve over time. In STM, this can be done by examining the effects of the year covariate on topical prevalence. Figure 11 plots the expected topic proportions against the year of publication for the first eight topics over the one hundred studied years. For each topic, we estimate the LOESS curve summarizing the data points along with the traditional 95% C.I. Thus, the graph shows the relative prevalence of each topic over time.

An eye-balling review of the graph reveals some interesting trends. For example, some topics such as “financial risk and return on investment” and “corporate social responsibility” show a declining time trend, indicating that academic research focusing on such topics was more likely to be published early on and less so recently. On the other hand, other topics such as “Emerging (East) Asian nations” and “global mergers and acquisitions” show an increasing trend, indicating that more papers were published recently. Other topics such as “moral and ethical issues” show temporal stability in terms of the prevalence of publications. Kang et al. (Forthcoming, p. 16) noted that “temporal stability may be a double-edged sword as on one hand, it reflects a field’s maturation and the existence of a committed group of scholars invested in this academic domain. On the other, it may hamper a disciplinary field from progress.” Once again, we can see that the topic of “systems and

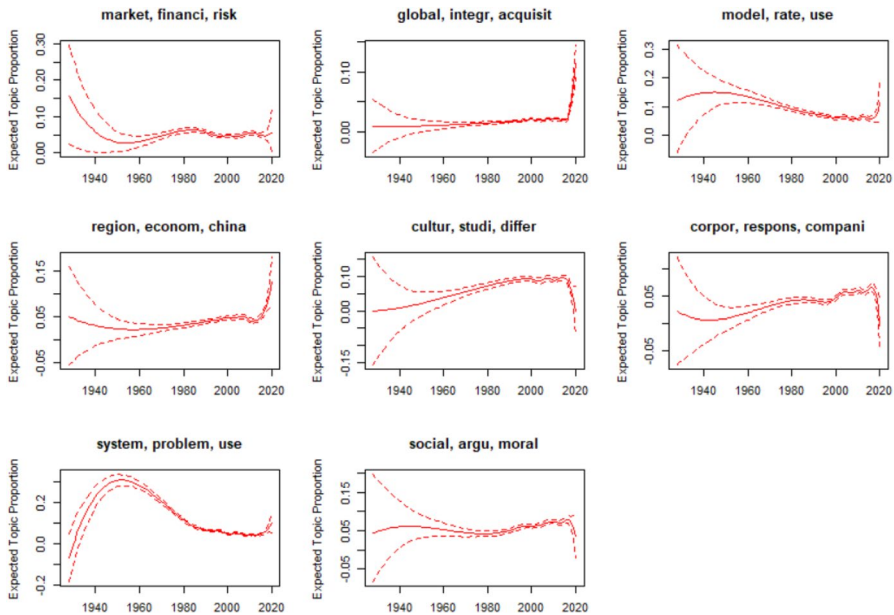


Fig. 11 Topic dynamics of selected international management research over time

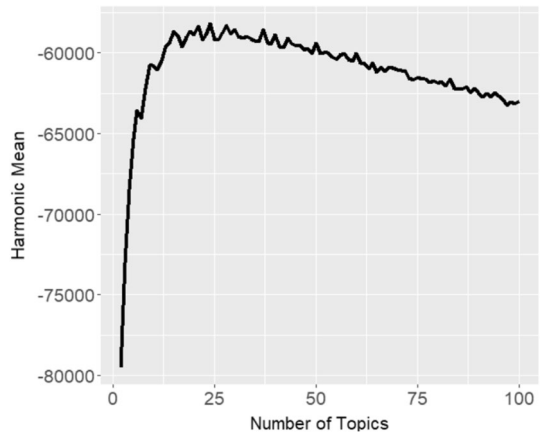
mathematical modeling” has reached its peak during the 1960s and early 1970s, confirming the “cliometric revolution” thesis. Although this “revolution” is marked by the publication of the article by Conrad and Meyer (1958), Cioni et al. (2020, p. 1) showed that it “took quite a long time to fully display its effects.”

Most of such trends are intuitive. For example, the burst in publications dealing with emerging Asian nations, including China, India and Japan over the last two decades is in line with the fact that international management scholars “have progressively focused their attention on emerging economies” (Pisani 2009, p. 214). Jack (2007, p. 364) noted that the success of Japanese corporations in the 1970s and 1980s was “a key impetus for the cross-cultural management literature and the culturist perspective more generally.” Research investigating emerging economies includes countries such as China (Li and Tsui 2002; Peng et al. 2001), Turkey (Demirbag et al. 1995) and Russia (Lawrance and Vlachoutsicos 1993).

4.5 LDA

To test the robustness of our results, we also apply the LDA technique to check whether our results hold. In fact, STM builds on LDA topic modeling algorithm and both were applied in several research settings (Han and Cho 2022). Following Sbalchiero and Eder (2020), the optimal number of topics was determined using the log-likelihood method. Figure 12 shows that the optimal number of topics determined by the algorithm is around the same number determined by the STM. Figure 13 shows the probabilities of the top terms per topic across the 18 topics. Because almost all the terms dealt with in each topic seem somehow related, no topics were classified as “junk” (Nikolenko et al. 2017). However, some topics were hard to interpret as is the case with topic 2. From the graph, we see

Fig. 12 The optimal number of topics (LDA)



that there are striking similarities between the LDA and the STM results. For example, the LDA topic 7 is similar to the STM topic 18 as both deal with “language and communication,” the LDA topic 11 is similar to the STM topic 1 as both deal with “financial risk and return on investment” and the LDA topic 11 is similar to the STM topic 11 as both deal with “legal settlement of disputes”. The preponderance of topics related to financial and legal issues, and the heavy reliance on economic, marketing and management science literature, demonstrates the interdisciplinary nature of international management research.

5 Discussion, limitations and future research

In this paper, we use a novel approach to extract latent academic topics in international management research using STM. Based on our findings, three major conclusions can be drawn. First, international management research is vast and multidisciplinary as it is heavily influenced by several other disciplines, such as Economics, Organizational Theory and Strategic Management. For example, some of the extracted topics are based on economic theories such as topic 3 (modeling cycles and estimating inflation rates) and 15 (transaction cost and supplier efficiency), while topic 17 (strategic alliances and organizational learning) is founded on organizational theories, and topic 10 (expatriation, gender and team-building skills) draws upon human resource management with a strategic orientation. This result is in line with Acedo and Casillas (2005, p. 632) who argued that “we should be cautious when considering international management as an independent field, a more accurate description being that of an interdependent field.” Second, attention to academic topics within international management research has varied over time. While some topics such as “financial risk and return on investment” and “corporate social responsibility” showed a declining time trend, other topics such as “Emerging (East) Asian nations” and “global mergers and acquisitions” showed an increasing trend, yet other topics such as “moral and ethical issues” showed temporal stability in terms of the prevalence of publications. The various levels of topic evolution over time indicate that international management research is characterized by constant flux. Third, academic topics in international management tend to cluster differently. For example, some topics are rarely combined together in existing research, which might be an indication of a research gap that can be exploited in future

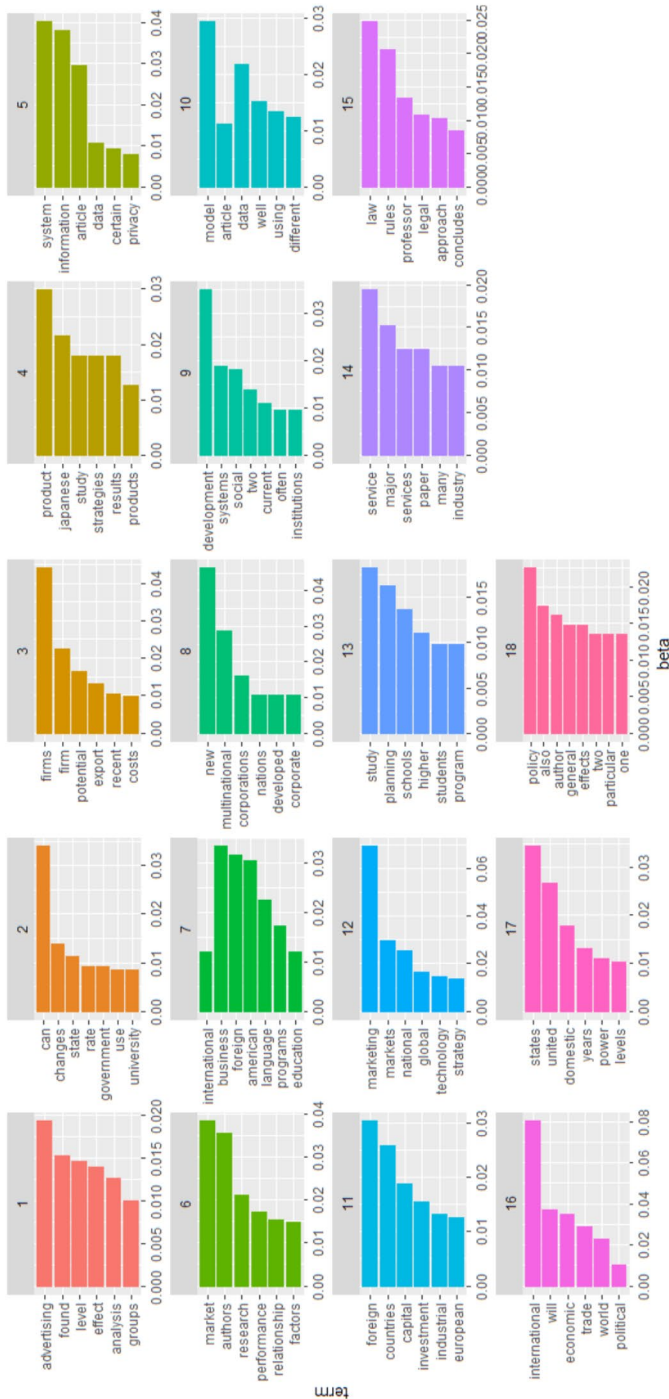


Fig. 13 Top terms in each international management research topic (LDA method)

research. For example, the topic related to culture (topic 5) and the topic related to (anti)-corruption policies (topic 9) are rarely investigated together, which implies that this might be a potential area of research in international management. Such rarely combined topics might also be an indication of the existence of a “tunnel vision” within the field of international management research, which might impede the theoretical development in the field. This conclusion is in line with Buckley (2002) who noted that international management research may be “running out of steam” and should search for a unifying “big question” for future inquiry.

Despite the major contribution of this study, it suffers from some limitations. For example, the results are based on a single database. Thus, results are inevitably affected by selection bias. Future research may combine the JSTOR with other databases such as WoS or Scopus to test whether our results hold. Second, we limited articles to the English language, which implies that some important research papers in other languages are not included in this analysis. Future research may combine such articles to test the robustness of our results. Third, we relied only on OCR scans of research articles, which might be prone to orthographical errors. However, Walker and Lund (2010) have shown that certain OCR random errors might be tolerated. Finally, topic modeling approaches cannot reveal “the reasons why specific topics are likely to be found together...nor can they give explanations about why specific topics did change over time” (Malaterre et al. Forthcoming). Thus, future research may combine topic modeling techniques with qualitative methods, which give the researcher the opportunity to reveal what Schwandt (2000, p. 190) calls “the fine-grained details” of research.

Funding The authors have not disclosed any funding.

Declarations

Conflict of interest The author declares no conflict of interest.

References

- Acedo, F., Casillas, J.: Current paradigms in the international management field: an author co-citation analysis. *Int. Bus. Rev.* **14**, 619–639 (2005)
- Adhikari, S., Dabbs, B.: Social network analysis in R: a software review. *J. Educ. Behav. Stat.* **43**, 225–253 (2018)
- Adler, N.: Understanding the ways of understanding: the ostrich and the trend. *Acad. Manag. Rev.* **8**, 226–232 (1983)
- Aleixandre, J., Aleixandre-Tudo, J., Bolanos-Pizarro, M., Aleixandre-Benavent, R.: Mapping the scientific research in organic farming: a bibliometric review. *Scientometrics* **105**, 295–309 (2015)
- Altaweel, M., Bone, C., Abrams, J.: Documents as data: a content analysis and topic modeling approach for analyzing responses to ecological disturbances. *Ecol. Inform.* **51**, 82–95 (2019)
- Ambrosino, A., Cedrini, M., Davis, J., Fiori, S., Guerzoni, M., Nuccio, M.: What topic modeling could reveal about the evolution of economics. *J. Econ. Methodol.* **25**, 329–348 (2018)
- Arino, A., de la Torre, J.: Learning from failure: toward an evolutionary model of collaborative ventures. *Organ. Sci.* **9**, 306–325 (1998)
- Bagozzi, B., Berliner, D.: The politics of scrutiny in human rights monitoring: evidence from structural topic models of US State Department Human Rights reports. *Polit. Sci. Res. Methods* **6**, 661–677 (2018)
- Bail, C., Brown, T., Mann, M.: Channeling hearts and minds: advocacy organizations, cognitive-emotional currents, and public conversation. *Am. Sociol. Rev.* **82**, 1188–1213 (2017)

- Bao, Y., Datta, A.: Simultaneously discovering and quantifying risk types from textual risk disclosures. *Manag. Sci.* **60**, 1371–1391 (2014)
- Barry, A., Valdez, D., Padon, A., Russell, A.: Alcohol advertising on Twitter: a topic model. *Am. J. Health Educ.* **49**, 256–263 (2018)
- Baumer, E., Mimno, D., Guha, E., Gay, G.: Comparing grounded theory and topic modeling: extreme divergence or unlikely convergence? *J. Am. Soc. Inf. Sci.* **68**, 1397–1410 (2017)
- Bellstam, G., Bhagat, S., Gookson, J.: A text-based analysis of corporate innovation. SSRN Working Paper No. 2803232 (2016)
- Bennett, R., Vijaygopal, R., Kottasz, R.: Who gives to food banks? A study of influences affecting donations to food banks by individuals. *J. Nonprofit Public Sect. Mark.* **(Forthcoming)**
- Blaydes, L., Grimmer, J., McQueen, A.: Mirrors for princes and sultans: advice on the art of governance in the medieval Christian and Islamic worlds. *J. Polit.* **80**, 1150–1167 (2018)
- Blei, D., Lafferty, J.: A correlated topic model of *Science*. *Ann. Appl. Stat.* **1**, 17–35 (2007)
- Boddeyn, J., Brewer, T.: International business political behavior: new theoretical directions. *Acad. Manag. Rev.* **19**, 119–143 (1994)
- Bohr, J., Dunlap, R.: Key topics in environmental sociology, 1990–2014: results from a computational text analysis. *Environ. Sociol.* **4**, 181–195 (2017)
- Borg, I., Groenen, B.: *Modern Multidimensional Scaling*. Springer, New York (1997)
- Bowen, H.: *Social Responsibilities of Businessmen*. Harper & Row, New York (1953)
- Buckley, P.: Is the international business research running out of steam? *J. Int. Bus. Stud.* **33**, 365–373 (2002)
- Caroll, A.: Corporate social responsibility: evolution of a definitional construct. *Bus. Soc.* **38**, 258–295 (1999)
- Celly, K., Spekman, R., Kamauff, J.: Technological uncertainty, buyer preferences and supplier assurances: an examination of Pacific Rim purchasing arrangements. *J. Int. Bus. Stud.* **30**, 287–316 (1999)
- Chen, X., Chen, J., Cheng, G., Gong, T.: Topics and trends in artificial intelligence assisted human brain research. *PLoS ONE* **15**, e0231192 (2020)
- Cheng, X., Shuai, C., Liu, J., Wang, J., Liu, Y., Li, W., Shuai, J.: Topic modeling of ecology, environment and poverty nexus: an integrated framework. *Agric. Ecosyst. Environ.* **267**, 1–14 (2018)
- Chikudate, N.: Generating reflexivity from partnership formation. *J. Appl. Behav. Sci.* **35**, 287–305 (1999)
- Child, J.: Culture, contingency and capitalism in the cross-national study of organization. In: Staw, B., Cummings, L. (eds.) *Research in organizational behavior*, pp. 303–356. JAI Press, New York (1981)
- Choi, H., Lee, W., Sohn, S.: Analyzing research trends in personal information privacy using topic modeling. *Comput. Secur.* **67**, 244–253 (2017)
- Chung, G., Rodriguez, M., Lanier, P., Gibbs, D.: Text-mining open-ended survey responses using structural topic modeling: a practical demonstration to understand parents' coping methods during the COVID-19 pandemic in Singapore. *J. Technol. Hum. Serv.* **(Forthcoming)**
- Cioni, M., Federico, G., Vasta, M.: The long-term evolution of economic history: evidence from the five field journals (1927–2017). *Cliometrica* **14**, 1–39 (2020)
- Cirillo, B., Tzabbar, D., Seo, D.: A bibliometric and topic modeling analysis of the structural divide in the multidisciplinary research on employee mobility. *Adv. Strateg. Manag.* **41**, 15–36 (2020)
- Coase, R.: The nature of the firm. *Economica* **4**, 386–405 (1937)
- Conrad, A., Meyer, J.: The economics of slavery in the Ante Bellum South. *J. Polit. Econ.* **66**, 95–130 (1958)
- Daenekindt, S., Huisman, J.: Mapping the scattered field of research on higher education: a correlated topic model of 17,000 articles, 1991–2018. *High. Educ.* **80**, 571–587 (2020)
- Demirbag, M., Mirza, H., Weir, D.: The dynamics of manufacturing joint ventures in Turkey and the role of industrial groups. *Manag. Int. Rev.* **35**, 36–51 (1995)
- Denny, M., Spirling, A.: Text preprocessing for unsupervised learning: why it matters, when it misleads, and what to do about it? *Polit. Anal.* **26**, 168–189 (2018)
- Ding, K., Choo, W., Ng, K., Ng, S.: Employing structural topic modeling to explore perceived service quality attributes in Airbnb accommodation. *Int. J. Hosp. Manag.* **91**, 102676 (2020)
- Dittmar, J., & Seabold, S. Media, markets and institutional change: the protestant reformation. Center for Economic Performance Discussion Paper. London, the UK (2015)
- Dybowski, T., Adammer, P.: The economic effects of U.S. presidential tax communication: evidence from a correlated topic model. *Eur. J. Polit. Econ.* **55**, 511–525 (2018)
- Dyer, J., Chu, W.: The determinants of trust in supplier-automaker relationships in the US, Japan, and Korea. *J. Int. Bus. Stud.* **31**, 259–285 (2000)
- Egri, C., Ralston, D.: Corporate responsibility: a review of international management research from 1998 to 2007. *J. Int. Manag.* **14**, 319–339 (2008)

- Fagnas, S., Hamilton, W., Espinosa, N., Miloff, A., Carlbring, P., Lindner, P.: What do users think about Virtual Reality relaxation applications? A mixed methods study of online user reviews using natural language processing. *Internet Interv.* **24**, 100370 (2021)
- Fan, D., Zhu, C., Huang, X., Kumar, V.: Mapping the terrain of international human resource management research: a bibliographic analysis. *J. World Bus.* **56**, 101185 (2021)
- Ferri, P., Lusiani, M., Pareschi, L.: Accounting for *Accounting History*: a topic modeling approach (1996–2015). *Account. Hist.* **23**, 173–205 (2018)
- Fey, C., Denison, D.: Organization culture and effectiveness: can American theories be applied in Russia? *Organ. Sci.* **14**, 686–706 (2003)
- Fino, E., Hanna-Khalil, B., Griffiths, M.: Exploring the public's perception of gambling addiction on Twitter during the Covid-19 pandemic: topic modeling and sentiment analysis. *J. Addict. Dis.* (**Forthcoming**)
- Flores, R.: Do anti-immigrant laws shape public sentiment? A study of Arizona's SB 1070 using Twitter data. *Am. J. Sociol.* **123**, 333–384 (2017)
- Franzosi, R., De Fazio, G., Vicari, S.: Ways of measuring agency: an application of quantitative narrative analysis to lynchings in Georgia (1875–1930). *Sociol. Methodol.* **42**, 1–42 (2012)
- Fresneda, J.E., Burnham, T.A., Hill, C.H.: Structural topic modeling segmentation: a segmentation method combining latent content and customer context. *J. Mark. Manag.* **37**, 792–812 (2021)
- Fruchterman, T., Reingold, E.: Graph drawing by force-directed placement. *Softw. Pract. Exp.* **21**, 1129–64 (1991)
- Gatti, C., Brooks, J., Nurre, S.: A historical analysis of the field of OR/MS using topic models (2015). [ArXiv:151005154](https://arxiv.org/abs/151005154).
- Gentzkow, M., Shapiro, J.: What drives media slant? Evidence from US daily newspapers. *Econometrica* **78**, 35–71 (2010)
- Goyal, N., Howlett, M.: “Measuring the mix” of policy responses to Covid-19: comparative policy analysis using topic modeling. *J. Comp. Policy Anal.* **23**, 250–261 (2021)
- Grajzl, P., Murrell, P.: Toward understanding 17th century English culture: a structural topic model of Francis Bacon's ideas. *J. Comp. Econ.* **47**, 111–135 (2019)
- Grajzl, P., Murrell, P.: Characterizing a legal-intellectual culture: Bacon, Coke and seventeenth-century England. *Cliometrica* **15**, 43–88 (2021)
- Greene, D., O'Callaghan, D., Cunningham, P.: How many topics? Stability analysis for topic models. *Lect. Notes Comput. Sci.* **8724**, 498–513 (2014)
- Griffiths, T., Steyvers, M.: Finding scientific topics. *Proc. Natl. Acad. Sci. USA* **101**, 5228–5235 (2004)
- Grimmer, J., Stewart, B.: Text as data: the promise and pitfalls of automatic content analysis. *Methods for political texts. Polit. Anal.* **21**, 267–297 (2013)
- Grothe-Hammer, M., Kohl, S.: The decline of organizational sociology? An empirical analysis of research trends in leading journals across half a century. *Curr. Sociol.* **68**, 419–442 (2020)
- Guo, Y., Barnes, S., Jia, Q.: Mining meaning from online ratings and reviews: tourist satisfaction analysis using Latent Dirichlet Allocation. *Tour. Manag.* **59**, 467–483 (2017)
- Gurcan, F., Cagiltay, N., Cagiltay, K.: Mapping human–computer interaction research themes and trends from its existence to today: a topic modeling-based review of past 60 years. *Int. J. Hum. Comput. Interact.* **37**, 267–280 (2021)
- Gutmann, M., Merchant, E., Roberts, E.: “Big data” in economic history. *J. Econ. Hist.* **78**, 268–299 (2018)
- Hagen, L.: Content analysis of e-petitions with topic modeling: how to train and evaluate LDA models? *Inf. Process. Manag.* (**Forthcoming**)
- Hajdinjak, S., Morris, M., Amos, T.: Migration policy framing in political discourse: evidence from Canada and the USA. In: Deutschmann, E., Lorenz, J., Nardin, L., Natalini, D., Wilhelm, A. (eds.) *Computational conflict research*. Springer, Berlin (2020)
- Hamel, G.: Competition for competence and inter-partner learning with international strategic alliances. *Strateg. Manag. J.* **12**, 83–104 (1991)
- Han, J., Cho, K.: Emerging industrial root technologies: a structural topic model- and topic matrix analysis-based approach. *Asian J. Technol. Innov.* **30**, 178–197 (2022)
- Han, C., Yang, M., Piterou, A.: Do news media and citizens have the same agenda on Covid-19? An empirical comparison of twitter posts. *Technol. Forecast. Soc. Change* **169**, 120849 (2021)
- Hansen, S., McMahon, M., Prat, A.: Transparency and deliberation within the FOMC: a computational linguistics approach. *Q. J. Econ.* **133**, 801–870 (2018)
- Hart, S.: Natural-resource-based view of the firm. *Acad. Manag. Rev.* **20**, 986–1014 (1995)
- Heiberger, R., Majo-Vasquez, S., Herrero, L., Nielsen, R., Esser, F.: Do not blame the media! The role of politicians and parties in fragmenting online political debate. *Int. J. Press Polit.* (**Forthcoming**)
- Heng, Y., Gao, Z., Jiang, Y., Chen, X.: Exploring hidden factors behind online food shopping from Amazon reviews: a topic mining approach. *J. Retail. Consum. Serv.* **42**, 161–168 (2018)

- Hofstede, G.: Motivation, leadership and organization: do American theories apply abroad? *Organ. Dyn.* **9**, 42–63 (1980)
- Hofstede, G.: *Cultures and Organizations: Software of the Mind*. McGraw Hill, London (1991)
- Holm, D., Eriksson, K., Johanson, J.: Creating value through mutual commitment to business network relationships. *Strateg. Manag. J.* **20**, 467–486 (1999)
- Hu, N., Zhang, T., Gao, B., Bose, I.: What do hotel customers complain about? Text analysis using structural topic model. *Tour. Manag.* **72**, 417–426 (2019)
- Huang, A., Lehavy, R., Zang, A., Zheng, R.: Analyst information discovery and interpretation roles: a topic modeling approach. University of Michigan School of Business Working Paper N0. 1229 (2016)
- Inkpen, A., Beamish, P.: Knowledge, bargaining, power, and the instability of international joint ventures. *Acad. Manag. Rev.* **22**, 177–202 (1997)
- Inkpen, A., Crossan, M.: Believing is seeing: joint ventures and organizational learning. *J. Manag. Stud.* **32**, 595–618 (1995)
- Jack, G.: International management and ethnography: what and why? *Ethnography* **8**, 361–372 (2007)
- Jelveh, Z., Kogut, B., Naido, S.: Political language in economics. Research Paper No. 14–57. Columbia Business School, New York, NY (2015)
- Jensen, J., Naido, S., Kaplan, E., et al.: Political polarization and the dynamics of political language: evidence from 130 years of partisan speech [with comments and discussion]. *Brook. Pap. Econ. Act.* **2**, 1–81 (2012)
- Jeyaraj, A., Zadeh, A.: Evolution of information systems research: insights from topic modeling. *Inf. Manag.* **57**, 103207 (2020)
- Jung, Y., Suh, Y.: Mining the voice of employees: a text mining approach to identifying and analyzing job satisfaction factors from online employee reviews. *Decis. Support Syst.* **123**, 113074 (2019)
- Kang, C., Baek, Y., Kim, E.: Half a century of NVSQ: thematic stability across years and editors. *Nonprofit Volunt. Sector Q.* (Forthcoming)
- Knudson, E., Dahlberg, S., Iversen, M., Johannesson, M., Nygaard, S.: How the public understands news media trust: an open-ended approach. *Journalism* (Forthcoming)
- Korfiatis, N., Stamolampros, P., Kourouthanassis, P., Sagiadinos, V.: Measuring service quality from unstructured data: a topic modeling application on airline passengers' online reviews. *Expert Syst. Appl.* **116**, 472–486 (2019)
- Kosnik, L.: What have economists been doing for the last 50 years? A text analysis of published academic research from 1960–2010. *Economics* **9**, 1–38 (2015)
- Kothari, T., Lahiri, S.: Yesterday, today and tomorrow: an overview of research publications in the Journal of International Management. *J. Int. Manag.* **18**, 102–110 (2012)
- Kushkowsky, J., Shrader, C., Anderson, M., White, R.: Information flows and topic modeling in corporate governance. *J. Doc.* **76**, 1313–1339 (2020)
- Kusters, A., Garrido, E.: Mining PIGS: a structural topic model analysis of Southern Europe based on the German newspaper *Die Zeit* (1946–2009). *J. Contemp. Eur. Stud.* **28**, 477–493 (2020)
- Kwon, K., Chadha, M., Wang, F.: Proximity and networked news public: structural topic modeling of global Twitter conversations about the 2017 Quebec Mosque shooting. *Int. J. Commun.* **13**, 2652–2675 (2019)
- Lang, C., Li, M., Zhao, L.: Understanding consumers' online fashion renting experience: a text-mining approach. *Sustain. Prod. Consum.* **21**, 132–144 (2020)
- Latzko-Toth, G., Bonneau, C., Millette, M.: Small data, thick data: thickening strategies for trace-based social media research. In: Sloan, L., Quan-Haase, A. (eds.) *The Sage handbook of social media research methods*, pp. 199–214. Sage, Los Angeles (2017)
- Lau, J., Newmann, D., Baldwin, T.: Machine reading tea leaves: automatically evaluating topic coherence and topic model quality. In: *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, Gothenburg, Sweden*, pp. 530–539 (2014)
- Lawrance, F., Vlachoutsicos, C.: Joint ventures in Russia: put the locals in charge. *Harv. Bus. Rev.* **71**, 44–51 (1993)
- Lee, J., Kolodge, K.: Exploring trust in self-driving vehicles through text analysis. *Hum. Factors* **62**, 260–277 (2020)
- Lemay, D., Baek, C., Doleck, T.: Comparison of learning analytics and educational data mining: a topic modeling approach. *Comput. Educ. Artif. Intell.* **2**, 100016 (2021)
- Li, X., Lei, L.: A bibliometric analysis of topic modelling studies (2000–2017). *J. Inf. Sci.* **47**, 161–175 (2021)
- Li, J., Tsui, A.: A citation analysis of management and organization research in the Chinese context: 1984–1999. *Asia Pac. J. Manag.* **19**, 87–107 (2002)

- Liu, X., Tvinnereim, E., Grimsrud, K., Lindhjem, H., Velle, L., Saure, H., Lee, H.: Explaining landscape preference heterogeneity using machine learning-based survey analysis. *Landsc. Res.* **46**, 417–434 (2021)
- Loree, D., Guisinger, S.: Policy and non-policy determinants of US equity foreign direct investment. *J. Int. Bus. Stud.* **26**, 281–299 (1995)
- Lu, J.: The evolving contributions in international strategic management research. *J. Int. Manag.* **9**, 193–213 (2003)
- Lucas, C., Nielsen, R., Roberts, M., Stewart, B., Storer, A., Tingly, D.: Computer-assisted text analysis for comparative politics. *Polit. Anal.* **23**, 254–277 (2015)
- Luo, J., Vu, H., Li, G., Law, R.: Topic modeling for theme park online reviews: analysis of Disneyland. *J. Travel Tour. Mark.* **37**, 272–285 (2020)
- Maerz, S., Schneider, C.: Comparing public communication in democracies and autocracies: automated text analyses of speeches by heads of government. *Qual. Quant.* **54**, 517–545 (2020)
- Makino, S., Delios, A.: Local knowledge transfer and performance: implications for alliance formulation in Asia. *J. Int. Bus. Stud.* **27**, 905–927 (1996)
- Malaterre, C., Lareau, F., Pulizzotto, D., St-Onge, J.: Eight journals over eight decades: a computational topic-modeling approach to contemporary philosophy of science. *Synthese* (**Forthcoming**)
- Martinez, Z., Toyne, B.: What is international management, and what is its domain? *J. Int. Manag.* **6**, 11–28 (2000)
- McFarland, D., Ramage, D., Chuang, J., Heer, J., Meaning, C., Jurafsky, D.: Differentiating language usage through topic models. *Poetics* **41**, 607–625 (2013)
- McInerney, F., Doherty, K., Bindoff, A., Robinson, A., Vickers, J.: How is palliative care understood in the context of dementia? Results from a massive open online course. *Palliat. Med.* **32**, 594–602 (2018)
- Melton, C., Olusanya, O., Ammar, N., Shaban-Nejad, A.: Public sentiment analysis and topic modeling regarding Covid-19 vaccines on the Reddit social media platform: a call to action for strengthening vaccine confidence. *J. Infect. Public Health* (**Forthcoming**)
- Minmo, D., Wallach, H., Talley, E., Leenders, M., McCallum, A.: Optimizing semantic coherence in topic models. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, USA, pp 262–272 (2011)
- Moro, S., Pires, G., Rita, P., Cortez, P.: A text mining and topic modeling perspective of ethnic marketing research. *J. Bus. Res.* **103**, 275–285 (2019)
- Mueller, H., Rauh, C.: Reading between the lines: prediction of political violence using newspaper text. *Am. Polit. Sci. Rev.* **112**, 358–375 (2019)
- Newman, D., Block, S.: Probabilistic topic decomposition of an eighteenth-century American newspaper. *J. Am. Soc. Inf. Sci. Technol.* **57**, 753–767 (2006)
- Nikolenko, S., Koltcov, S., Koltsova, O.: Topic modelling for qualitative studies. *J. Inf. Sci.* **43**, 88–102 (2017)
- Ovadek, M.: “Popular tribunes” and their agendas: topic modeling Slovak presidents’ speeches 1993–2020. *East Eur. Polit.* **37**, 214–238 (2021)
- Pandur, M., Dobsa, J., Kronegger, L.: Topic modeling in social sciences: case study of Web of Science. In: *Central European Conference on Intelligent and Information Systems* (2020)
- Park, S., Lim, Y., Park, H.: Comparing Twitter and YouTube networks in information diffusion: The case of the “Occupy Wall Street” movement. *Technol. Forecast. Soc. Change* **95**, 208–217 (2015)
- Park, E., Park, J., Hu, M.: Tourism demand forecasting with online news data mining. *Ann. Tour. Res.* **90**, 103273 (2021)
- Parkhe, A.: “Messy” research, methodological predispositions, and theory development in international joint ventures. *Acad. Manag. Rev.* **18**, 227–268 (1993)
- Peng, M., Lu, Y., Shenkar, O., Wang, D.: Treasures in the China house: a review of management and organizational research on Greater China. *J. Bus. Res.* **52**, 95–110 (2001)
- Pisani, N.: International management research: Investigating its recent diffusion in top management journals. *J. Manag.* **35**, 199–218 (2009)
- Ploessl, F., Just, T., Wehrheim, L.: Cyclicity of real estate-related trends: topic modeling and sentiment analysis on German real estate news. *J. Eur. Real Estate Res.* (**Forthcoming**)
- Priva, U., Austerweil, J.: Analyzing the history of *Cognition* using topic models. *Cognition* **135**, 4–9 (2015)
- Puschmann, C., Ausserhofer, J., Slerka, J.: Converging on a nativist core? Comparing issues on the Facebook pages of the Pegida movement and the Alternative for Germany. *Eur. J. Commun.* **35**, 230–248 (2020)
- Quinn, K., Monroe, B., Golaresi, M., Grespin, M., Radev, D.: How to analyze political attention with minimal assumptions and costs. *Am. J. Polit. Sci.* **54**, 209–228 (2010)

- R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna (2021)
- Rao, A., Schmidt, S.: A behavioral perspective on negotiating international alliances. *J. Int. Bus. Stud.* **29**, 665–694 (1998)
- Remmers, L.: International financial management: 35 years later- What has changed? *Int. Bus. Rev.* **13**, 155–180 (2004)
- Ricks, D., Toyne, B., Martinez, Z.: Recent developments in international management research. *J. Manag.* **16**, 219–253 (1990)
- Roberts, M., Stewart, B., Tingley, D., Airoidi, E.: The structural topic model and applied social science. Advances in neural information processing systems workshop on topic models: computation, application and evaluation. Lake Tahoe (2013)
- Roberts, M., Stewart, B., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S., Albertson, B., Rand, D.: Structural topic models for open-ended survey responses. *Am. J. Polit. Sci.* **58**, 1064–1082 (2014)
- Roberts, M., Stewart, B., Airoidi, E.: A model of text for experimentation in the social sciences. *J. Am. Stat. Assoc.* **111**, 988–1003 (2016)
- Rodriguez, M., Storer, H.: A computational social science perspective on qualitative data exploration: using topic models for the descriptive analysis of social media data. *J. Technol. Hum. Serv.* **38**, 54–86 (2020)
- Roose, H., Roose, W., Daenekindt, S.: Trends in contemporary art discourse: using topic models to analyze 25 years of professional art criticism. *Cult. Sociol.* **12**, 303–324 (2018)
- Sainju, B., Hartwell, C., Edwards, J.: Job satisfaction and employee turnover determinants in Fortune 50 companies: insights from employee reviews from Indeed.com. *Decis. Support Syst.* **148**, 113582 (2021)
- Sanchez-Franco, M., Arenas-Marquez, F., Alonso-Dos-Santos, M.: Using structural topic modeling to predict users' sentiments towards intelligent personal agents: an application for Amazon's Echo and Google Home. *J. Retail. Consum. Serv.* **63**, 102658 (2021)
- Savin, I., Drews, S., van den Bergh, J.: Free association of citizens and scientists with economic and green growth: a computational-linguistics analysis. *Ecol. Econ.* **180**, 106878 (2021)
- Sbalchiero, S., Eder, M.: Topic modeling, long texts and the best number of topics: some problems and solutions. *Qual. Quant.* **54**, 1095–1108 (2020)
- Schmiedel, T., Muller, O., vom Brocke, J.: Application example on organizational culture. *Organ. Methods Res.* **22**, 941–968 (2018)
- Schwandt, T.: Three epistemological stances for qualitative inquiry: interpretivism, hermeneutics, and social constructionism. In: Denzin, N.K., Lincoln, Y.S. (eds.) *Handbook of qualitative research*, vol. 2, pp. 189–213. Sage Publications, Thousand Oaks (2000)
- Sharma, A., Rana, N., Nunkoo, R.: Fifty years of information management research: a conceptual structure analysis using structural topic modeling. *Int. J. Inf. Manag.* **58**, 102316 (2021)
- Sharpe, W.: Capital asset prices: a theory of market equilibrium under condition of risk. *J. Finance* **19**, 425–442 (1964)
- Shrader, R., Oviatt, B., McDougall, P.: How new ventures exploit trade-offs among international risk factors: lessons for the accelerated internationalization of the 21st century. *Acad. Manag. J.* **43**, 1227–1247 (2000)
- Shrivastava, P.: *Greening Business: Profiting the Corporation and the Environment*. Thompson Executive Press, Cincinnati (1996)
- Simonin, B.: Ambiguity and process of knowledge transfer in strategic alliances. *Strateg. Manag. J.* **20**, 595–624 (1999)
- Srinivas, S., Rajendran, S.: Topic-based knowledge mining of online student reviews for strategic planning in universities. *Comput. Ind. Eng.* **128**, 974–984 (2019)
- Stamolampros, P., Korfiatis, N., Chalvatzis, K., Buhalis, D.: Job satisfaction and employee turnover determinants in high contact service: insights from employees' online reviews. *Tour. Manag.* **75**, 130–147 (2019)
- Starik, M., Rands, G.: Weaving an integrated web: multilevel and multisystem perspectives of ecologically sustainable organizations. *Acad. Manag. Rev.* **20**, 908–935 (1996)
- Stening, B., Skubik, D.: Do international management researchers need a code of ethics? *Manag. Int. Rev.* **47**, 103–126 (2007)
- Stewart, B., Zhukov, Y.: Use of force and civil-military relations in Russia: an automated content analysis. *Small Wars Insur.* **20**, 319–343 (2009)
- Sun, W., Wang, X.: A raciolinguistic analysis of the neoliberal promotion of dual language education in a new Latinex South state. *Discourse Stud. Cult. Polit. Educ.* (**Forthcoming**)

- Sun, L., Yin, Y.: Discovering themes and trends in transportation research using topic modeling. *Transp. Res. Part C* **77**, 49–66 (2017)
- Symitsi, E., Stamolampros, P., Daskalakis, G., Korfiatis, N.: The informational value of employee online reviews. *Eur. J. Oper. Res.* **288**, 605–619 (2021)
- Taddy, M.: On estimation and selection for topic models. *Artif. Intell. Stat.* **22**, 1184–1193 (2012)
- Terman, R.: Islamophobia and media portrayals of Muslim women: a computational text analysis of US news coverage. *Int. Stud. Q.* **61**, 489–502 (2017)
- Tornberg, A., Tornberg, P.: Combining CDA and topic modeling: analyzing discursive connections between Islamophobia and anti-feminism on an online forum. *Discourse Soc.* **27**, 401–422 (2016)
- Tvinnereim, E., Liu, X., Jamelske, E.: Public perceptions of air pollution and climate change: different manifestations, similar causes and concerns. *Clim. Change* **140**, 399–412 (2017)
- Walker, D., Lund, W.: Evaluating models of latent document semantics in the presence of OCR errors. In: *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 240–250 (2010)
- Warwick, D.: The politics and ethics of cross-cultural research. In: Triandis, H., Lambert, W. (eds.) *Handbook of cross-cultural psychology 1: Vol.: Perspectives*, pp. 319–371. Allyn and Bacon, Boston (1980)
- Wehrheim, L.: Economic history goes digital: topic modeling the *Journal of Economic History*. *Cliometrica* **13**, 83–125 (2019)
- Werner, S.: Recent developments in international management research: a review of 20 top management journals. *J. Manag.* **28**, 277–305 (2002)
- Williamson, O.: *Markets and Hierarchies: Analysis and Antitrust Implications*. Free Press, New York (1975)
- Williamson, O.: Transaction cost economics: the governance of contractual relations. *J. Law Econ.* **22**, 223–261 (1979)
- Williamson, O.: *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Free Press, New York (1985)
- Winson-Geideman, K., Evangelopoulos, N.: Topics in real estate research, 1973–2010: a latent semantic analysis. *J. Real Estate Lit.* **21**, 59–76 (2013)
- Yan, A., Gray, B.: Bargaining power, management control, and performance in the United States-China joint ventures: a comparative case study. *Acad. Manag. J.* **37**, 1478–1517 (1994)
- Zaheer, A., Zaheer, S.: Catching the wave: alertness, responsiveness, and market influence in global electronic networks. *Manag. Sci.* **43**, 1493–1509 (1997)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.