Available online at www.sciencedirect.com
ScienceDirect
Journal of Interactive Marketing 40 (2017) 52 – 72
JOURNAL OF INTERACTIVE MARKETING
www.elsevier.com/locate/intmar

Popular Research Topics in Marketing Journals, 1995–2014

Yung-Jan Cho [a] & Pei-Wen Fu [b] & Chi-Cheng Wu [b,*]

[a] *College of Management, National Sun Yat-sen University, Taiwan*
[b] *Department of Business Management, National Sun Yat-sen University, Taiwan*

Available online 3 October 2017

Abstract

During the past two decades, the focus of marketing has moved from the tactics of persuasion to the strategies of value cocreation. After moving toward cognitive science and corporate strategies in the early 2000s, marketing research returned to its traditional domains of consumer psychologies and customer management. While conscientious consumers are gradually restraining themselves from selfish indulgence, marketers have refocused on a new set of values that encompass mental, experiential, and societal well-being. In this regard, we adopt an unprecedented approach by incorporating topic modeling with social network analysis. The results show that, in terms of topic heterogeneity, the most impactful journals are the most diverse, whereas each runner-up has a unique focus. Among the journals, we detect two major co-authorship communities, and among the topics, we detect three. Further, we find that the communities of the most cited papers are composed of heterogeneous clusters of similar topics. The pivots within, and the bridges between, these communities are also reported. In the spirit of collaborative research, our topic model and network analysis are shared via online collaboration and visualization platforms that readers can use to explore our models interactively and to download the dataset for further studies.

Keywords: Topic model; Text mining; Data visualization; Reproducible and collaborative research

Introduction

In the same way that they predict the trends in markets, data analytics may pinpoint the evolving trends within academia. By applying text-mining technology to journal articles, marketing researchers have identified the historical trends of popular research topics and thereby predicted the future direction of marketing science. The current study extends this research thread in three ways. First, while prior studies have each focused on a single journal, we cover all articles published in 25 journals (see Table 1). Second, leveraging recent advancements in data science, we discover more topics and trends in greater detail. Third, in the spirit of collaborative research, we upload our data to the Internet, visualize the results, and make them interactively accessible via online collaboration platforms.

The universal corpus enables us to compare the topical focuses across journals and capture the overarching trend. Advanced techniques lead to many previously undocumented topics and insights. By combining topic modeling and social network analysis, we also examine the coauthorship structure across journals, topics, and the most cited papers. With these improvements, we aim to augment the aforementioned studies with a broader scope, greater detail, and better presentation.

Background and Literature Review

Natural language processing is a rapidly evolving technology that has been widely adopted in academia. In this regard, scholars have recently begun to adopt text analysis technology to study the topical history of marketing research. In three consecutive years, MS (*Marketing Science*: For conciseness we assign an abbreviated code to each of the 25 journals studied here—please see Table 1 for a complete list of the codes), JMR (the *Journal of Marketing Research*), and JCR (the *Journal of Consumer Research*) have each published a paper that elucidates the "topical history" of their articles. The authors have respectively studied 30, 50, and 40 years of historical text

* Corresponding author.
E-mail address: chicheng@mail.nsysu.edu.tw (C.-C. Wu).

https://doi.org/10.1016/j.intmar.2017.06.003
1094-9968/

Table 1
The scope of analysis (25 SSCI journals, 1995–2014).

Journal name	Code	Listed since*	No. papers	No. cites	No. cites per paper	Impact factor**
European Journal of Marketing	EJM	2005	714	6,631	9.3	1.09
Industrial Marketing Management	IMM	1995	1,550	29,169	18.8	1.93
International Journal of Research in Marketing	IJRM	1997	482	10,526	21.8	1.83
International Marketing Review	IMR	1999	425	5,946	14.0	1.59
Journal of the Academy of Marketing Science	JAMS	1997	608	29,513	48.5	3.74
Journal of Advertising	JA	1995	576	13,332	23.1	2.29
Journal of Advertising Research	JAR	1995	741	10,705	14.4	0.99
Journal of Business Research	JBR	1995	2,908	50,214	17.3	2.13
Journal of Consumer Affairs	JCA	1995	395	5,479	13.9	1.05
Journal of Consumer Psychology	JCP	2000	576	12,119	21.0	2.01
Journal of Consumer Research	JCR	1995	1,079	48,619	45.1	3.19
Journal of International Marketing	JINTL	1995	391	8,725	22.3	3.25
Journal of Interactive Marketing	JIM	2007	160	2,967	18.5	3.26
Journal of Macromarketing	JMM	2008	188	1,096	5.8	1.43
Journal of Marketing	JM	1995	744	68,091	91.5	3.89
Journal of Marketing Research	JMR	1995	990	47,228	47.7	3.11
Journal of Product Innovation Management	JPIM	1995	804	24,649	30.7	2.09
Journal of Public Policy & Marketing	JPPM	1995	436	7,084	16.2	1.15
Journal of Retailing	JR	1995	547	22,188	40.6	2.01
Journal of Services Marketing	JSM	2009	275	1,277	4.6	1.02
Journal of Service Research	JSR	2004	271	6,744	24.9	2.46
Marketing Letters	ML	2001	396	4,660	11.8	1.51
Marketing Science	MS	1995	833	28,569	34.3	1.65
Psychology & Marketing	PM	1995	1,024	18,110	17.7	1.37
Quantitative Marketing & Economics	QME	2006	130	1,112	8.6	0.85
			17,243		24.9	2.03

1. For data integrity, all of the original data in this study, including the paper abstracts, the numbers of citations and the impact factors, were extracted from a single source—the *Web of Science*™ database maintained by *Thomson Reuters*™.
2. The 'listed since*' column above indicated the years when the journals were listed in SSCI. Usually, journals were listed several years after they were founded.
3. Impact Factor** was acquired from *Journal Citation Report*® 2015 in the *World of Science*® database.

extracted from MS, JMR, and JCR. While these studies all employed text analysis technology, the topic modeling methods differ from each other.

Mela, Roos, and Deng (2013) studied a corpus of 1,085 papers published in MS since 1982. Rather than finding topics from the text, the authors directly analyzed keywords and the most frequent words used in the papers' abstracts. They discovered that most of the popular words matched the traditional scope, the 4Ps and 3Cs, of marketing. Popular methods and subjects have evolved over time, from 'logit' and 'consumer choice' in the 1980s, to 'econometrics' and 'brand' in the 1990s, and to 'Bayesian,' 'Internet,' and 'customer relationship management (CRM)' in the 2000s. The authors observed that, while the introduction rate of new keywords increased over time, the survival rate of such words declined. Consequently, the authors posited that marketing science may have evolved from youth to maturity.

Huber, Kamakura, and Mela (2014) applied ontological learning, a dictionary-based topic modeling method (see Wong, Liu, and Bennamoun (2012) for a review), to 2,531 papers published in JMR and discovered 44 topics from the corpus. The authors observed that during the past 50 years, the popularity of 'information processing' has continued to rise; that of 'advertising,' 'choice model,' and 'measurement' has been falling; and that of 'consumer behavior' seems to have revived from a low point at the turn of the century.

Leveraging the current mainstream topic modeling method, latent Dirichlet allocation (LDA), Wang et al. (2015) analyzed 2,031 articles published in JCR since 1974. Although this study employed a powerful analytic method, it only acquired 16 topics from the corpus. The authors found that the popularity of 'methodology' and 'family decision' has continuously decreased; that of 'advertising,' 'memory and persuasion,' and 'customer satisfaction' has declined from its climax; and that of 'consumer culture' has continuously increased during the entire period (from 1974 to 2014). Based on their observation, the authors predicted that 'consumer culture,' 'social identity and influence,' and 'emotional decision' may be among the most promising research topics in the near future.

Enhancements and Distinctive Findings

The major objective of this research is to extend the breadth and depth of the aforementioned studies. Regarding scope, each of the aforementioned papers is dedicated to a single journal, and each of these journals has a specific area of interest. In contrast, we cover the whole realm of marketing with a corpus taken from 25 marketing journals (selected from the Social Science Citation Index (SSCI): see Table 1). In terms of corpus size, this study is at least 20 times broader than any of the prior works.

Such a universal corpus enables us to align the overarching trend of marketing to the dramatic societal changes triggered by the Internet and globalization since the 1990s. When the role models for business excellence changed from those of Toyota, General Electronics, and Walmart to Amazon, Starbucks, and Apple, the focus of marketing moved from product quality, customer satisfaction, and channel coordination to brand difference, service dominance, and customer experience. During the past two decades, researchers' focus has moved from marketing tactics to value cocreation strategies. While conscientious consumers are gradually restraining themselves from selfish indulgence, marketers have refocused on mental, experiential, and societal well-being. From a piecewise perspective, some of our findings may resemble our predecessors'. However, none of the journal-specific studies has covered the entire realm of marketing and captured the overarching trend, as aforementioned.

Text analysis technology has evolved swiftly. Consequently, this study utilizes several updated technologies that were not adopted in the prior studies. By using the techniques of multi-gram tokenization and noun-phrase extraction, our vocabulary is much more expressive than that of prior works. For example, our model discerns 'customer satisfaction (CS),' 'CRM,' 'customer lifetime value (CLV),' and 'customer experience (CX)'; thereby, we observe how the implication of 'customer' has evolved over time. Moreover, cluster computing platforms and distributed topic modeling algorithms help to extract more topics from a larger corpus. Compared with Wang et al. (2015), which defined 16 topics from 235 single words, our model identifies 100 topics from more than 2,000 phrases. From advertisement to branding strategies, behavioral, cognitive to social psychologies, product, channel, and service dominance to customer centrality, and globalization and regulation to industry-specific topics, it appears that a large number of topics are required to cover the widely spread domain of marketing research.

In order to facilitate our analysis, we further categorize the 100 'topics' into 13 "topic groups". Hereafter, to be discernible, we enclose the names of 'topics' and "topic groups" in single and double quotes respectively. Such a hierarchical arrangement helps to manage the topical details in an organized way. For example, "customer and service" not only manifests the evolution of customer centrality but also reveals its close relationship with service management. While prior studies separately identified 'choice,' 'memory,' 'persuasion,' and 'emotion' as individual topics, we treat "cognitive" and "behavioral psychologies" as two large domains. The waxing and waning of subordinate topics, when systematically tracked, imply that today's marketing researchers are more interested in serving consumers' mental interests than exploiting their cognitive flaws. Within the hierarchy of topics, we also identify several previously overlooked research areas. For example, "value network" addresses the growing interest in value proposition, business models, and value cocreation networks. The aggregated share of industry-specific topics, such as 'food,' 'health care' and 'movies,' is larger than that of the analytic methods; however, industry-specific topics have been entirely neglected in prior studies.

Based on the outcomes of topic modeling, we visualize the evolving path of marketing research on the topical space (Fig. 5, the right panel). The path basically reflects the hitherto paradigm shift from product quality and channel management, via market orientation and brand differentiation, to customer centrality and service experience. Despite its recent prominence, digital marketing seems to fall off the evolving path. Arguably, it might be situated at the fringe of marketing and serve as a bridge that connects to management information system (MIS) and computer science.

This study adopts an unprecedented approach by examining topical heterogeneity across marketing journals. Based on topical composition, most of the journals specialize in "strategy and value," "product, channel, and services," and/or "psychology and advertising" (Fig. 5, the left panel). Besides these traditional domains, a few relatively 'young' journals occupy certain 'niche' topical areas. The *Journal of Interactive Marketing* (JIM) specializes in digital marketing, the *Journal of Service Research* (JSR) in service management, and the *Journal of International Marketing* (JINTL) in globalization strategies. The most impactful journals, namely the *Journal of Marketing* (JM) and the *Journal of the Academy of Marketing Science* (JAMS), have the largest topical diversities, whereas each runner-up journal has a unique specialty area. Intriguingly, the runners-up happen to be the aforementioned young and niche journals. Their impact factors are second only to JM and JAMS (see Fig. 6).

This study extends its predecessors' social network analyses in two directions. First, we identify previously unreported network characteristics. We show that the small-world phenomenon is exhibited in the authors' and papers' networks, but the long-tail phenomenon is only presented in the former. By comparing the edge-weight distributions, we find that marketing researchers prefer to cooperate with new coauthors than to remain within static cliques. Across academic disciplines, we find that connectivity[1] in marketing's research network (71.6%) is higher than that of library and information science (20%), but lower than those of biomedicine (92.6%) and physics (85.4%).

Second, by incorporating the results of topic modeling, we examine the structure of coauthorship among the journals, topics (groups), and papers. Coauthorship among journals (see Fig. 9a) is composed of two major communities. The first community gathers around an axis of MS–JMR–JCR, which roughly reflects the general domain of marketing research. The second is centered on the *Journal of Business Research* (JBR) and connects to journals with specific interests such as international marketing, product marketing, and advertisement. The coauthorship network among the topics is composed of three communities (see Fig. 9b). These specialize respectively in strategies, analytics, and psychologies, which, arguably, may reflect the major knowledge domains of marketing researchers.

Topic modeling and social network analyses both generate valuable information that cannot be fully expressed in static tables and charts. In order to achieve better presentation and accessibility, we uploaded our model and most of the associated data to an online collaboration platform (GitHub) and presented them via web-based visualization tools (Tableau); thus, readers can browse the topic

[1] In terms of the proportional coverage of the giant component within the social network.

model in dynamic web pages, interactively query the citation database, customize the display for their own interests, and download the full set of data for further research.

The following sections are arranged as follows. The Data Source and Preprocessing section describes the data source and data preprocessing. The Topic Modeling section explains the processes of topic modeling, naming, and grouping. In the Trend Analysis section, we analyze the trends and discuss their implications. Journal comparison and social network analysis are presented in the Citation and Impact and Social Network Analysis (SNA) sections. Limitations and ideas for future studies are then reported in the Discussion, Limitations, and Further Studies section.

Data Source and Preprocessing

First, we identified 25 marketing journals from the SSCI. Then, we extracted all articles published in these journals from 1995 to 2014 from the *Web of Science*™ database. After removing book reviews and editorial articles, there were 17,249 research papers left in our corpus.

Using the "tm" package (the text-mining toolkit in R, see Meyer, Hornik, and Feinerer (2008)), we transferred words in the abstracts into plain, lowercase text, and removed punctuation and numbers. Then, we purged the general English stop words (the "SMART" stop words suggested by Manning, Raghavan, and Schütze (2008)) and a set of context-specific stop words, including study, paper, research, journal, finding, editor, article, issue, publish, discussion, manuscript, literature, and their plural forms.

After stemming (with the algorithm suggested by Porter (1980)), we applied bi-gram tokenizing to the corpus. This tokenization produced 680,255 terms. We removed the terms that occurred less than five times, calculated the term frequency–inverse document frequency (tf–idf) scores (see Chowdhury 2010), and then truncated the lowest quantile (as suggested by Hornik and Grün (2011)). Finally, we censored the terms with excessive prevalence: 'brand,' 'advertise,' and 'price,'[2] and removed the documents that were left with fewer than six terms. This process led to a matrix of 17,243 documents and 44,917 terms (814,399 tokens).

Topic Modeling

After exploring the currently available methods,[3] tools,[4] and algorithms,[5] it appears that R's 'lda' package, which fits LDA topic models with collapsed Gibbs sampling, produces the best result[6] for our corpus. Following the work of Wang et al. (2015), we started with a rather flat prior, $\alpha=\eta=0.02$, which implies even distributions of topics within the documents (Blei, Ng, and Jordan 2003). In the tuning process, we found that smaller values for α and η tend to produce more meaningful topics. After scanning $0.002 \le \alpha, \eta \le 0.2$, we set (α, η) at (0.01,0.0075). With regard to the number of topics, after considering model fitness and interpretability, we stopped at 100 topics, where the number of ambiguous topics starts to increase. Details of the parameter tuning procedure and model selection criteria are reported in Appendix C.

Topic Naming and Topic Dictionary

After being identified by the model, the topics still had to be manually named. Adequate naming is crucial for the interpretability of topic models. However, naming topics may be as difficult as "reading tea leaves" (Chang et al. 2009). Traditionally, topics were named after the most likely terms, ranked by $p(w|t)$, which represents the probability of words for a topic. However, with such an approach, frequent terms tend to occupy the naming space and make the topics barely distinguishable. Following the work of Sievert and Shirley (2014), we ranked the terms within each topic by their *relevance*, defined as:

$$r(w,t|\lambda) = \lambda \log[p(w|t)] + (1-\lambda) \log\left[\frac{p(w|t)}{p(w)}\right] \tag{1}$$

where λ divides the weight between the conditional probability of a term and its lift. As suggested by Sievert and Shirley (2014), we found that setting $\lambda=0.6$ helps to identify the most relevant and salient terms.

The 100 topics are alphabetically sorted in a topic dictionary (see Appendix A) where each topic name is followed by its defining terms listed in descending relevance. For example, the topic 'ad creativity' is defined by the terms creativity, ad agent, ad creativity, award, creativity process, and so on. Further, 'ad media' is defined by television, media, commercial, television ads, audience, program, broadcast, and similar. Note that because of its excessive prevalence, the word "advertising" was censored during preprocessing. However, with multi-gram tokens, our model not only captures the significance of advertising; it also classifies its relevant papers into three distinguishable topics: 'ad creativity,' 'ad media,' and '(print) advertisements'. The same applies to "brand," which is divided into 'brand equity,' 'brand extension,' and 'brand image,' and "price," which is embedded in 'willingness to pay (WTP) and price (discrimination),' 'promotion (price elasticity),' 'stock price,' and 'price (perception) effects'.

Aside from dividing such significant terms as 'brand' and 'advertising' into separate topics, the LDA model also joins

[2] Words of excessive prevalence would skew the topic model. Because of bi-gram tokenizing, phrases such as "brand equity", "advertising media", and "price elasticity" remain in the document-term matrix and remain significant in the topic model, as will be explained later.

[3] Latent semantics indexing (LSI) (Deerwester et al. 1990) and LDA (Blei, Ng, and Jordan 2003).

[4] R's lsa package (Wild, Rstem, and Wild 2009), R's lda package (Chang 2015), R's topicmodels package (Hornik and Grün 2011), and Apache Spark MLlib LDA package.

[5] Expectation maximization (EM), variational inference, and collapsed Gibbs sampling.

[6] Judged by term relevancy, topic meaningfulness, and the log likelihood of the model.

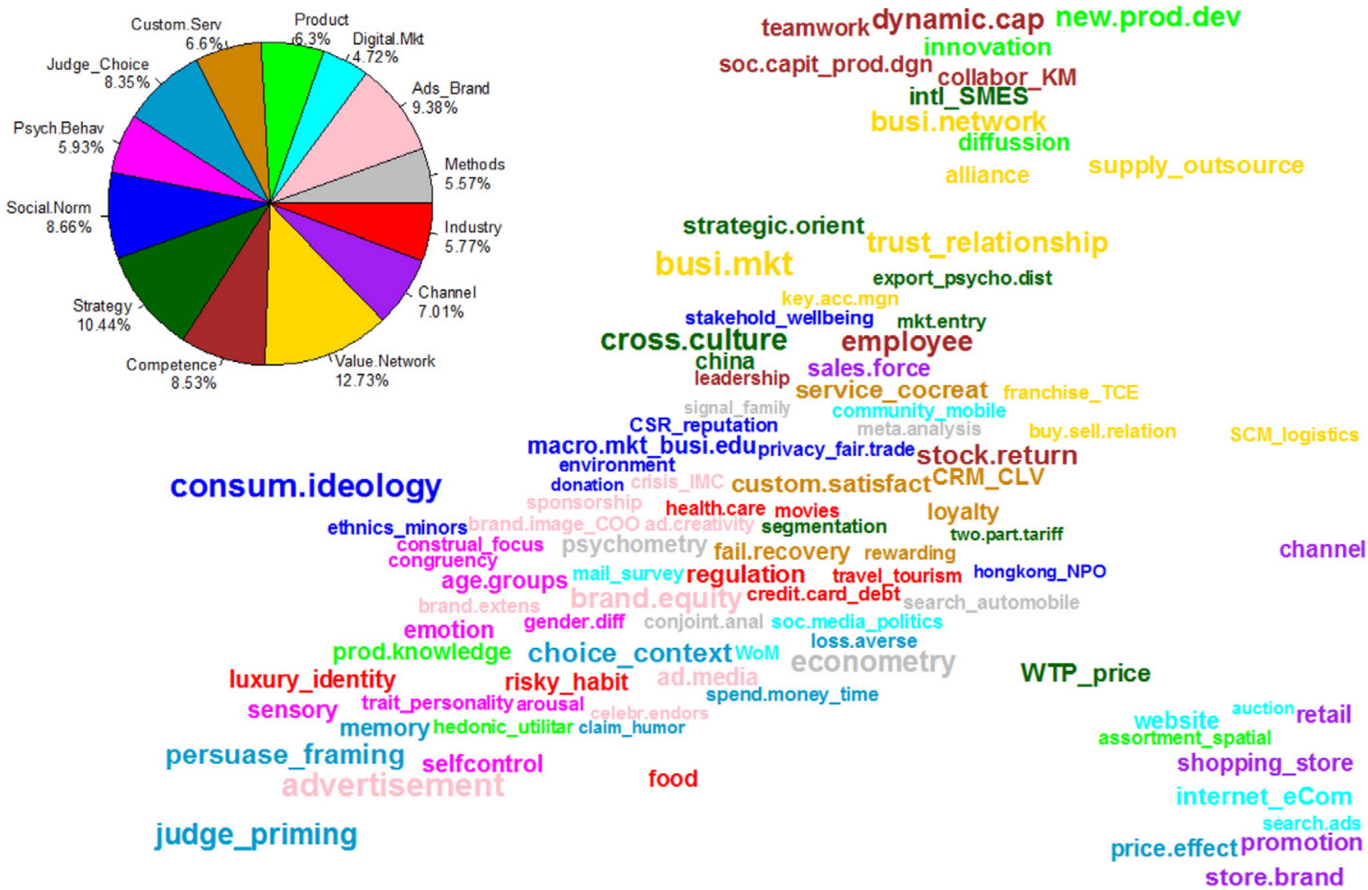

Fig. 1. The 100 topics in 13 groups.

interrelated terms to make ‘compounded topics,’[7] such as ‘assortment and spatial,’ ‘brand image and COO,’ and ‘choice and context’. The seemingly unrelated elements in most of these compounds are actually connected by some subtle links. For example, ‘spatial’ models are widely adopted in product ‘assortment’; every country (‘COO’) has its unique ‘brand image’; and studies of ‘context’ effect often adopt ‘choice’ models. These compounded topics also help to reveal latent relationships among popular research topics. For example, the performance of knowledge management (‘KM’) depends on ‘collaboration’; the perceived value of ‘luxury’ goods is associated with social ‘identity’; consumers often ‘search’ for information before they buy ‘automobiles’; ‘risk’ perception is related to the consumption of ‘habitual’ substances; and so on.

Topic Groups and the Topic Word Cloud

Based on their underlying knowledge domains, we categorized the 100 topics into 13 topic groups. As shown in Fig. 1, the 13 groups are indicated by distinctive colors. The 100 topics are presented as a word cloud, in which the font sizes and colors represent the topics’ relative prevalence and the groups to which they belong respectively. The words’ positions are derived from the posterior[8]; thus, proximities between the words imply the similarity of topics within the model. As illustrated in Fig. 1, words of the same color tend to stay together. It implies that our topic model, which determines the words’ locations, is consistent with our domain knowledge, which determines the colors.

Model Visualization

Besides allocating topics to articles, the topic model also informs us how terms are distributed among topics, and how topics are connected via terms. However, such a delicate relationship between terms and topics cannot be properly expressed in either the dictionary (Appendix A) or the word cloud (Fig. 1). Inspired by Sievert and Shirley (2014), we uploaded the entire model to GitHub and made it accessible via a dynamic web page (see this URL).[9] Readers can explore the topic model for their own particular interests using an interactive, intuitive interface, as displayed in Fig. 2.

[7] All of the compounded topics are listed at the end of the dictionary (Appendix A). The reasons that bind the compounding elements are also provided in the list.

[8] By applying a principal coordinate analysis (Gower 1966) to the terms’ probability matrix for the topics.

[9] http://bl.ocks.org/tonychuo/raw/2483473d1f924dfa0702/#topic=0&lambda=0.6&term=.

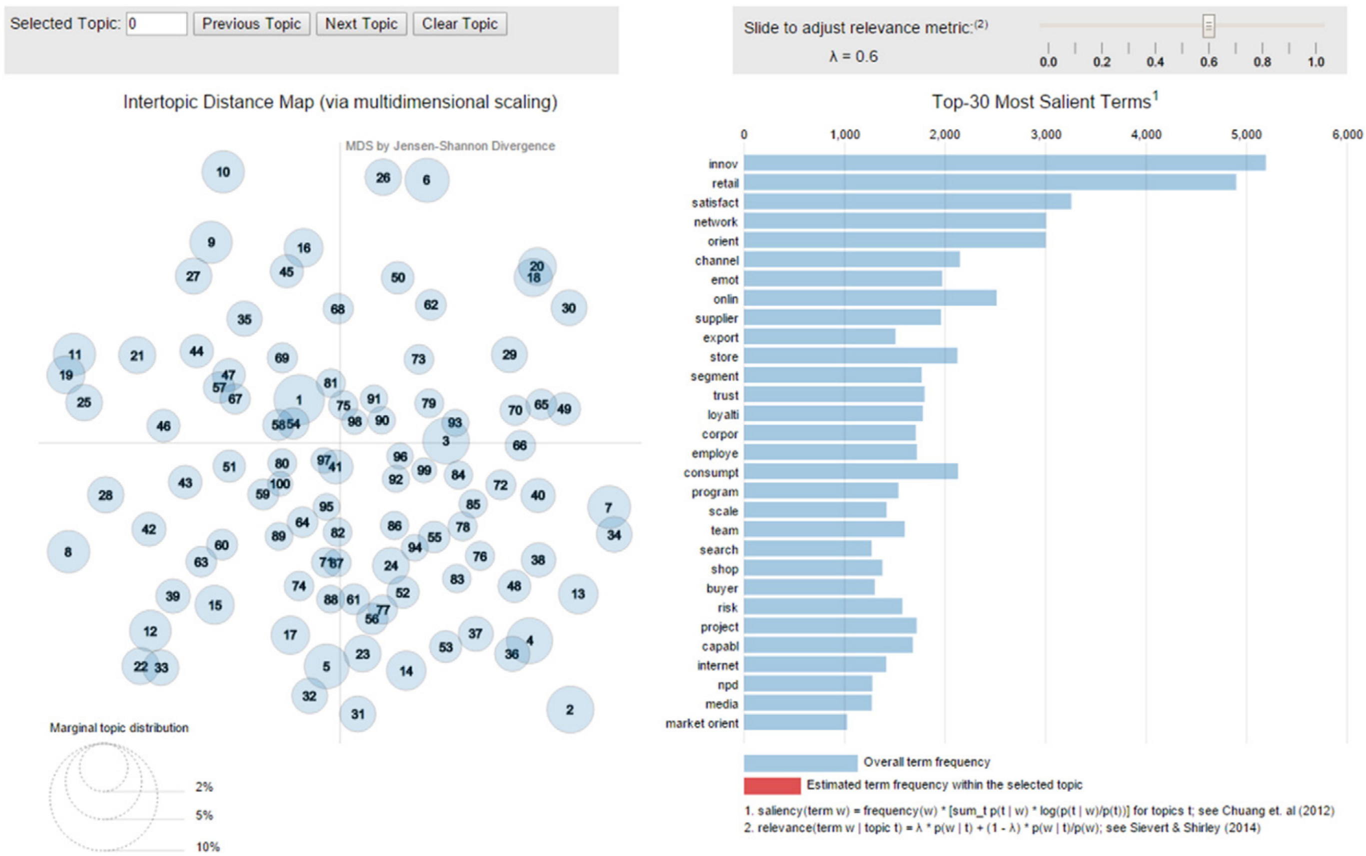

Fig. 2. Topic model browser. URL: http://bl.ocks.org/tonychuo/raw/2483473d1f924dfa0702/#topic=0&lambda=0.6&term=.

Trend Analysis

Reflecting the fact that a paper may span several topics, LDA employs a hierarchical Bayesian model in which topics are allocated to each article in multinomial distributions. From the posterior, we can derive the probabilities of every topic within every document. Following the convention of LDA (Blei, Ng, and Jordan 2003), hereafter we *ascribe* the most likely topic to each document. Given a (set of) document(s), the *weight* of a topic is calculated as the sum of its probabilities within the document(s).

The Overall Trend of Marketing Research

The 13 topic groups jointly delineate the general scope of marketing research. The relative weights of these groups are presented in the pie chart in Fig. 1. The largest group is “value network” (11.27%); it addresses the topics pertaining to value chain, business model, and business networks, which, jointly, may reflect marketers’ growing interest in platform- and network-based business models. The smallest group is “customer and service” (5.84%), which encompasses customer management (CRM/CLV/customer value management (CVM)) and service (experience) related topics. That said, judging relative importance by comparing group weights could be misleading. Because the groups consist of different numbers of topics, examining the temporal trends between and within the groups should be more sensible than directly comparing their weights.

The overall trends of the 13 topic groups are presented in Fig. 3. By applying Loess regression to each group’s annual weights, we can observe how each group’s popularity changes over time. As illustrated in the upper left of Fig. 3, “advertising and branding” was the leading topic group 20 years ago; however, its popularity has been falling. Further, when we trace its subordinate topics (shown in the lower left panel of Fig. 4), we find that ‘advertisement’ and ‘ad media’ were the most popular topics, but their popularity has continuously declined. Although ‘brand equity’ has been on the rise, its rising momentum is not sufficient to offset the falls of advertising. The overall popularity of this former leader has thus regressed to the mean.

“Value network” grew and took the lead in the late 1990s. “Strategy,” which addresses pricing, segmentation, market entry, and globalization strategies, and “judgment and choice,” which concerns the cognitive processes of consumption decisions, also exhibit concave trend lines. Although these three topic groups still hold the lead, the margin by which they lead has eroded during the first decade of the twenty-first century.

In the upper right of Fig. 3, we can identify three fast-rising groups. “Social norms” addresses societal and humanitarian issues such as corporate social responsibility (CSR), the environment, sustainability, fair trade, and poverty. “Psychology and behavior” investigates consumers’ behavior and psychologies. “Competence” concerns corporates’ core capabilities such as dynamic capability,

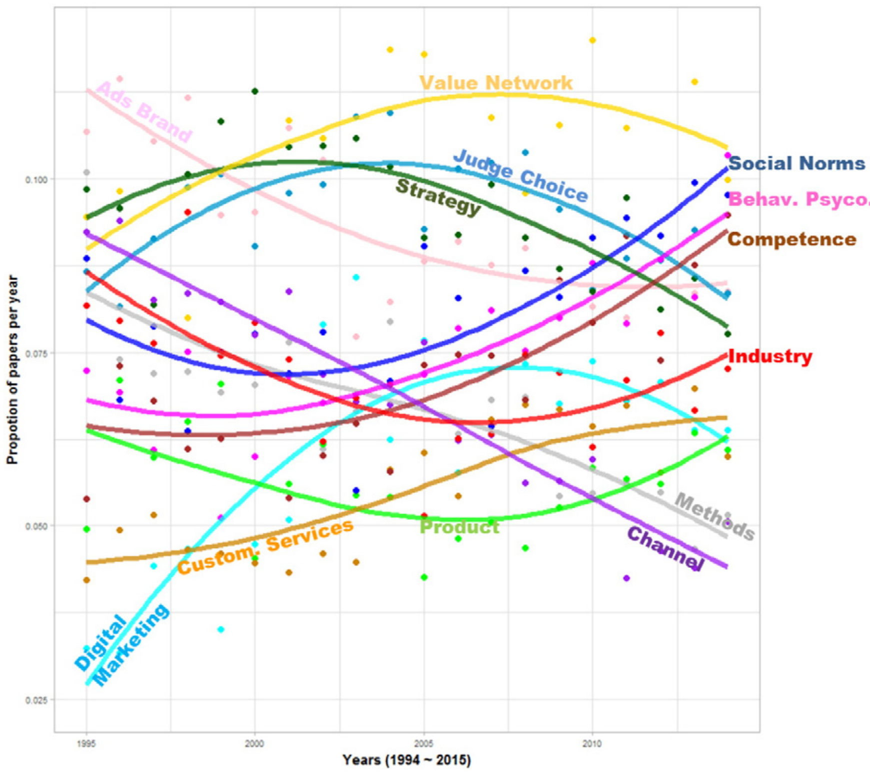

Fig. 3. Trends of the topic groups.

knowledge management, teamwork, and leadership. These three groups have taken the second to fourth places since 2012 and, at this pace, “social norms” may take the lead by 2020.

Two falling topic groups are noted in the lower right of Fig. 3. “Methods” encompasses econometrics, psychometrics, and meta-analysis, and “channel” covers topics related to retail,

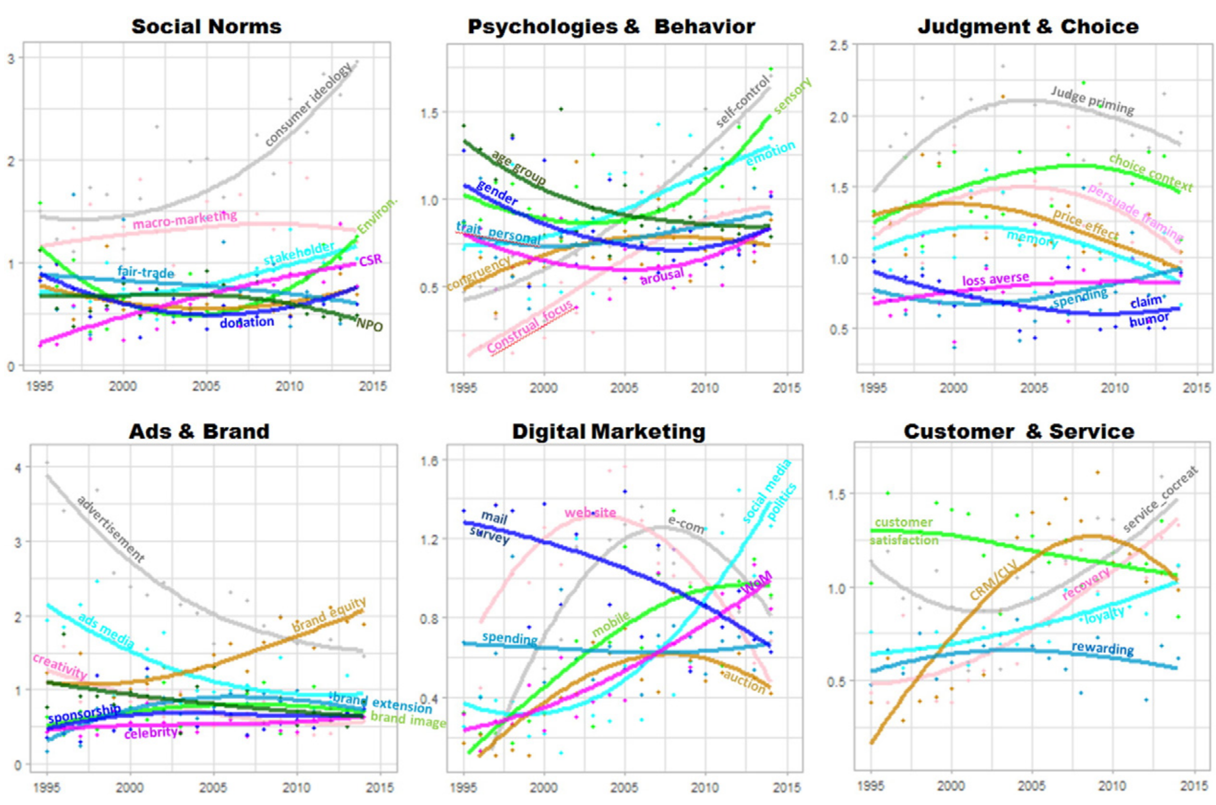

Fig. 4. Trend within each topic groups.

store branding, and sales forces. Collectively, these trends imply that marketing researchers have redirected their efforts from methods to applications, and from such marketing techniques as advertisement and channel distribution to value proposition, value networks, and consumer psychologies.

The Trend Within Each Topic Group

We can observe the trend within each topic group by plotting the weights of its subordinate topics over time. The trend within the fastest-rising topic group, "social norms," is shown in the upper-left panel of Fig. 4. Most of the rising momentum in this group comes from 'consumer ideology'. Narratives regarding marketing-resistance and anti-consumption have become popular over time. Such a trend may reflect the postmodern consumers' introspective intention to conscience consumption and their urge for existential authenticity.

Under the "psychology and behavior" group (in the upper-middle panel of Fig. 4), we see the decline of two demographics, 'age group' and 'gender,' and the rise of three psychological topics, 'self-control,' 'sensory,' and 'emotion'. We also observe the emergence of the self-related constructs, '(self) construal and (regulatory) focus'. Regarding "judgment and choice," most of us should have witnessed the era of the "predictably irrational" (Ariely 2008) in which researchers intensively explored how priming and framing techniques may apply to marketing practices. As indicated in the upper-right panel of Fig. 4, judgment, priming, persuasion, and framing climaxed around 2005 and gradually declined after then. Such a phenomenon may imply that, while applying psychologies to marketing, the ultimate objective is to increase consumers' mental well-being rather than exploit their cognitive flaws.

Compatible with our experience, the lower-left panel of Fig. 4 reveals how researchers' interests changed from 'advertisement' to 'branding'. Contrastingly, some of the "digital marketing" topics (the lower-middle of Fig. 4) are not as popular in academia as we have witnessed in the marketplace. Such buzzwords as 'mobile marketing,' 'social marketing,' and 'word of mouth' do not obtain the same level of popularity as we have expected. The trend of "customer and service" topics is illustrated in the lower right of Fig. 4. As 'customer satisfaction' slowed down, some of its momentum has transferred to 'loyalty'. In contrast to the rise and fall of database marketing ('CRM/CLV'), the revival of the 'experience cocreation' topic is noteworthy. The increasing popularity of CX (see Lemon and Verhoef (2016) for a review) shows how 'consumer' and 'service' coincide at 'experience,' which according to Vargo and Lusch (2008) is the essence of service and can only be cocreated by consumers and producers.

By applying first-order regression to the annual weights of topics, we identified the fastest rising and falling topics. The fastest rising topics are 'business network,' 'consumer ideology,' 'luxury–identity,' 'self-control,' and 'innovation'. Meanwhile, the topics of 'advertisement,' 'promotion,' 'regulation,' 'econometrics,' and 'ad media' are the fastest falling. Here again, we witness the rise of value proposition and the fall of marketing tactics. Such a trend suggests how the functional role of marketing has changed during the past 20 years. Rather than an attracting and persuading technique (via advertisement and promotion), nowadays marketing has been repositioned as a value-creation activity. While conscientious consumers are gradually abandoning selfish indulgence, marketers are refocusing on a new set of value propositions. These propositions encompass not only hedonistic and individualistic goods, but also mental and societal well-being.

The Evolving Path of Topical Trends

On the right of Fig. 5, we mark the evolving path of marketing research on the correspondence space of topic groups. We find that researchers' interest has moved from products and advertisement, channels and methodologies, value and choice, customers and service, to consumer behaviors and social norms. In the past two decades, the trend of marketing research basically reflects the shifting paradigm from product quality, customer satisfaction, market orientation, channel power, and customer relationship to brand differentiation, service dominance, and customer centrality. Along the evolving path, we find an intriguing turnaround in 2005: Rather than evolving toward the seemingly promising realm of digital marketing, the trend line bends backward to the traditional marketing domain, namely customer services and consumer psychologies. Owing to its mixed knowledge domain, digital marketing may be situated at the fringe of marketing, serving as a bridge that connects to the disjointed disciplines of MIS and computer sciences.

Citation and Impact

Journal Comparison

The association between the topic groups and journals is presented as a correspondence map on the left of Fig. 5. It appears that most journals and topics align on the vertical central line of the map. The journals and topics are roughly grouped by their distinctive knowledge domains, namely "value network and strategy," "product, channel, and analytics," and "psychologies and advertisement".

By incorporating citation, impact factor,[10] and topical information, we further compare the 25 journals in Fig. 6, where the sizes of the floating pies indicate impact factors and the sectors within the pies represent the composition of topics. The vertical axis is the average number of citations per paper. Because it takes time to accumulate citations, summing up the 20-year data would strongly favor the senior journals. Thus, the average citation is calculated with data later than 2007. The horizontal axis represents the topical diversity of the journals. Inspired by Leydesdorff and Rafols (2011), we define the topical diversity of a journal j as:

$$d_j = 1/\sum_t \left(\frac{w_{j,t}}{W_j}\right)^2 \quad \text{where} \quad W_j = \sum_t w_{j,t} \tag{2}$$

where $w_{j,t}$ is the weight of topic group t in journal j.

[10] Gathered from http://impactfactor.cn on December 2016.

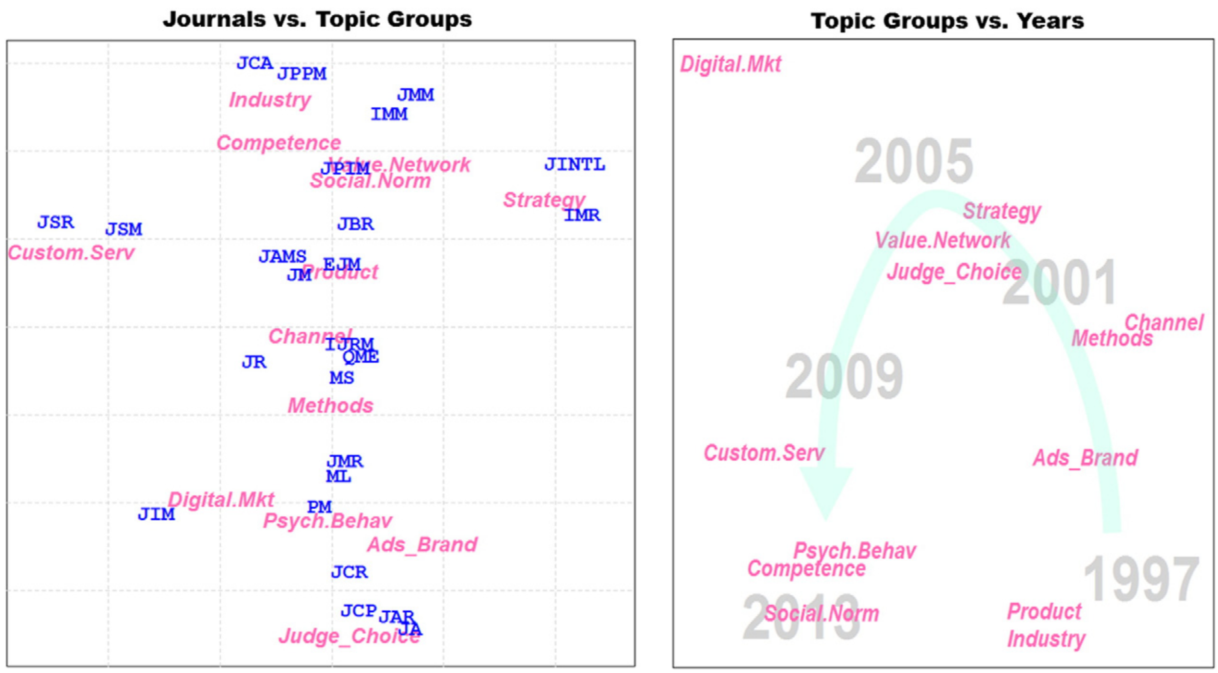

Fig. 5. Topic-journal association and topic evolving path.

In Fig. 6, the larger pies stay on the top with most of them scattered along the diagonal of the positive slope. While both the impact factor and topical diversity positively correlate to the average number of citations, the correlation between the impact factor and topical diversity is not significant. The most impactful journals, JM and JAMS, both have a 'balanced' topic composition,

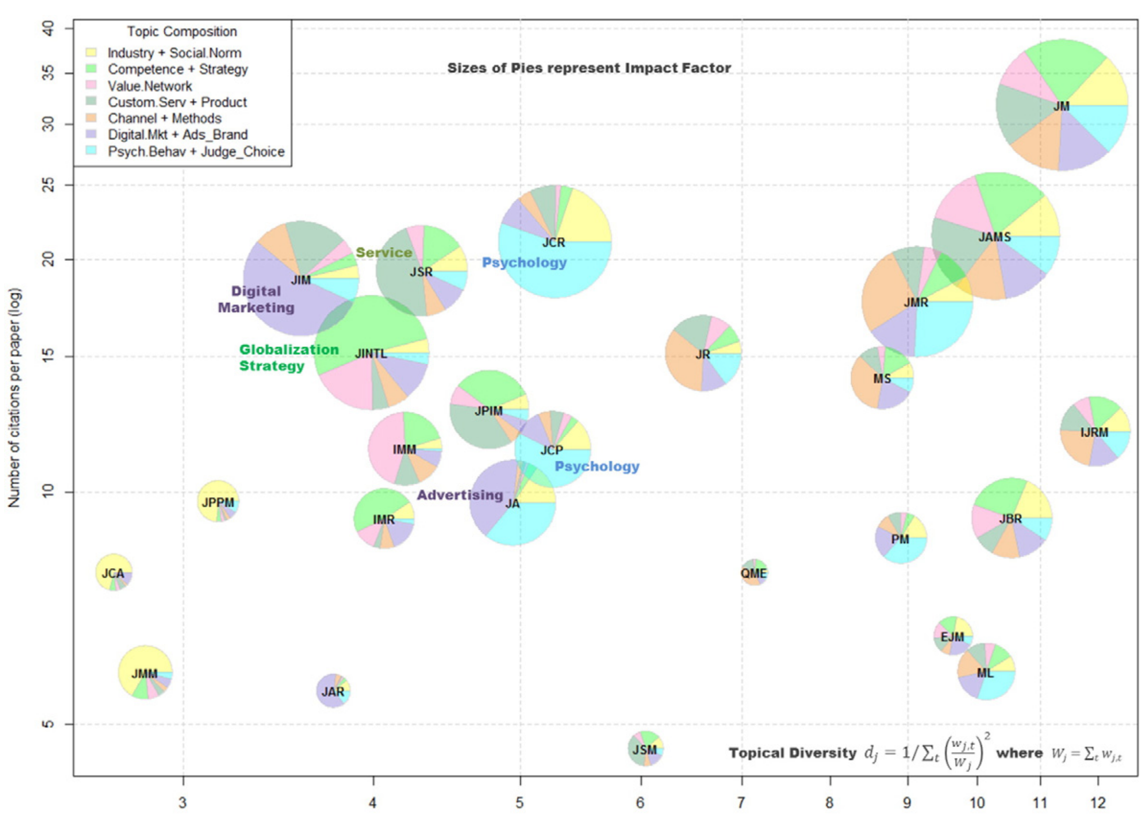

Fig. 6. Comparison among journals (2007–2014).

whereas each of the runner-up journals has unique interest. In descending order of impact factor, JCR, JINTL, JIM, *Journal of Advertising Research* (JAR), and JSR specialize in topics related to consumer psychologies, globalization strategy, digital marketing, advertisement, and services respectively. On the upper-left above the diagonal, we can identify a group of three specialized journals, JIM, JSR, and JINTL. If we refer back to the correspondence map (on the left of Fig. 5), we find that these relatively young yet impactful journals each occupy a niche topical area: JIM for digital marketing, JSR for services, and JINTL for globalization. Arguably, their outstanding performance may be attributed to the increasing popularity of their specialized topic areas.

Topic Heat Map and Online Citation Map

In order to illustrate the volume of citations at the topic level, we prepared a citation heap map (see Web Appendix 2), where the average citations for the 100 topics and 25 journals are juxtaposed. We find that the most frequently cited topics are 'CSR,' 'psychometrics,' 'customer satisfaction,' and a group of digital marketing topics that include 'online community and mobile marketing,' 'web sites,' and 'Internet and e-commerce.' As we can observe in Web Appendix 1, JIM particularly specializes in these digital marketing topics. This specialization may explain why JIM maintains the highest average citation rate among all the junior journals that began after 2005.

In order to make our results accessible, we consolidated the citation data with the topic model and visualized the result via the *Tableau Public*™ publication platform (see www.tableau.com). With this URL, readers can interactively query the database by topics, topic groups, and/or journals, and see how the selected papers are distributed by time on a dynamic web page.[11] The instructions for using the citation map are provided in Web Appendix 3.

Social Network Analysis (SNA)

The corpus of 17,243 papers includes 16,410 unique authors. On average, each paper has 2.47 coauthors, and each author contributes to 2.6 papers. The distributions of papers' coauthorship and authors' productivity are presented in Fig. 7a and b respectively. As shown, the latter exhibits a much longer tail than the former. At the low end, 60.3% of the authors only contribute to one paper; at the high end, 614 authors (3.91%) are involved in more than 10 papers.

Characteristics of the Social Network

The paper–author association is transformed into two undirected graphs, namely the *authors' network* and the *papers' network*, in which the edge weights represent the numbers of cooperating papers (common authors) between two authors (papers). The major characteristics of these two networks are presented in Fig. 8.

As Fig. 8 illustrates, the densities in both networks are low. Only 1 of 1,000 pairs of papers has a common author, while it takes 4,000 pairs of authors to find one example of coauthorship. Despite their low densities, the transitivities in both networks' large components are high, with 62.2% in the papers' network and 25.3% in the authors' network. Such transitivities imply that the chance of coauthorship with one's coauthors' coauthor is about one in four (25.3%). Moreover, two papers are more likely (62.2%) to have a common author than not if they both have coauthorship with a third paper.

From the average path lengths (5.28 for papers and 6.16 for authors), we can observe the "small world" property in both networks. On average, only three papers exist between two connected papers, and four authors between two connected authors. As noted on the right of Fig. 8, the authors' degree distribution exhibits a linear pattern (after a log–log transformation), but the papers' does not. This finding indicates that the authors' network is in a condition of power distribution, while the papers' network is exponential. The "long tail," which implies "the rich get richer" phenomenon, only occurs in the former.

As shown on the left of Fig. 8, the weight distribution of the papers' network decays rapidly. Among the 17,243 papers, only 16 pairs have more than five authors in common. This result indicates that sizeable, long-term coauthorship is rare. Marketing researchers do not stay in static cliques; they are willing to cooperate with new partners in order to achieve diversity and novelty.

Our analysis shows that 71.6% of marketing authors are interconnected within the giant component—the largest interconnected subgraph. Such a result is consistent with Goldenberg et al. (2010) which reported a 69% giant component in the social network among marketing researchers. The giant component is one of the key metrics for measuring the integrity and connectivity of a social network. The aforementioned result implies that a considerable portion of marketing authors (~30%) is located in some small 'islands' that are completely isolated from the main community.

In order to ensure completeness, we also compared our results with coauthorship studies in other disciplines. Benchmarking to the coauthorship network in library and information science, where the network size and the average number of papers per author are similar to ours, we found that the giant component only covers 20% of the authors (Goldenberg et al. 2010). According to Newman (2001), the sizes of the giant components in the researcher networks of biomedicine (92.6%) and physics (85.4%) are larger than that of marketing (71.6%). However, these networks are much larger than ours (1,520,251 and 52,909 authors compared with 16,410 authors) and their authors are more productive on average (6.4 and 5.1 papers per author compared with 2.6 papers per author).

The Structure of Coauthorship Network

By aggregating the edge weights and applying community detection algorithm (Blondel et al. 2008), we further examined the coauthorship structure at the journal, topic group, and topic levels. The structure of the journal network is presented in Fig. 9a, where the node sizes and the edge widths represent the

[11] http://public.tableau.com/views/JN20/Dashboard1.

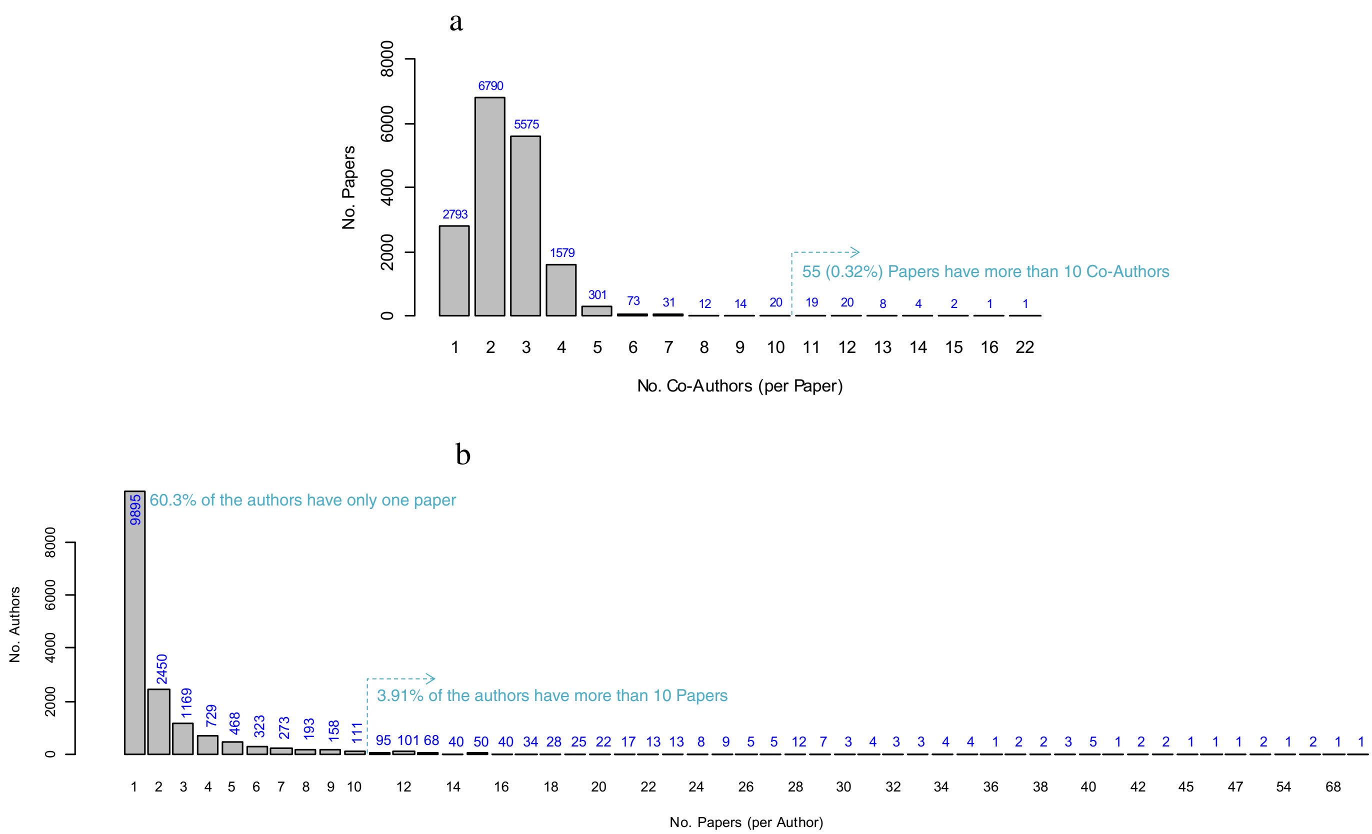

Fig. 7. a. Distribution of the papers' co-authorship. b. Distribution of the authors' productivity.

number of papers in the journals and the strength of the coauthorship between them.[12] Within the journal networks, we identify two communities. On the left-hand side, the thick belt of MS–JMR–JCR manifests the general scope of marketing research, with analytics at one (MS) end and psychologies at the other (JCP). Arguably, the prestige of JMR may contribute to its pivotal position in this community. Sitting at the center of the right-hand community, JBR connects to several specialized journals such as *Industrial Marketing Management* (IMM) and *Psychology & Marketing* (PM) for product and industrial marketing, the *Journal of Advertising* (JA) and JAR for advertisements, and JINTL and the *International Marketing Review* (IMR) for international marketing. Owing to JBR's large amount of published work, its authors also bridge the two communities by connecting to journals on the left-hand side.

Coauthorship among the topic groups is presented in Fig. 9b. The tripartite network structure basically reflects the three major knowledge domains in marketing; namely, strategies (the lower-left), methods (the upper community), and psychologies (the lower-right). The modularity (the ratio of within versus between community edge strength, see Blondel et al. (2008)) is low and there is no obvious center within the communities. Every node in the upper community has multiple external edges. This finding implies that "methods" is a common knowledge domain shared by researchers of strategies and psychologies.

After filtering out eigenvector centrality (the lower 50%) and edge weight (the lower 18%), the coauthorship network of the most connected topics is presented in Fig. 9c. The structure at topic level is similar to, but not fully consistent with, that at the level of topic groups. Analytics and channel-related topics are hosted in the upper-right community. Behavioral and cognitive psychologies gather with advertisement in the upper-left. The lower community is quite heterogeneous: It encompasses a company's capabilities, business relationships, brand, culture, ideology, and even psychometrics. Between the upper two communities, the bridging edges among econometrics, price, and promotion, and among choice, judgment, and advertisement are notable. The strong ties among them indicate that the analytical and psychological communities are mainly connected via these topics.

In order to ensure completeness, we also conducted a structural analysis on the network of the most cited papers. Of the 884 papers that have more than 100 citations, we detected 11 communities. When we examined their topical composition, we found that each of these communities is composed of heterogeneous clusters that consist of similar topics. In other words, papers of similar topics form clusters, and clusters of heterogeneous topics form communities. Connectivity among the

[12] Any individual is counted as a coauthor between two journals if and only if he or she ever (co)published at least one paper in each of the two journals. The strength of coauthorship between two journals is defined as the number of coauthors between them. The edge widths are proportionate to the strength of coauthorship transformed by the logarithm. Actually, coauthorship exists between almost every pair of journals. In order to present a clear structure, we have censored the edges whose weights are lower than 15% of the maximum.

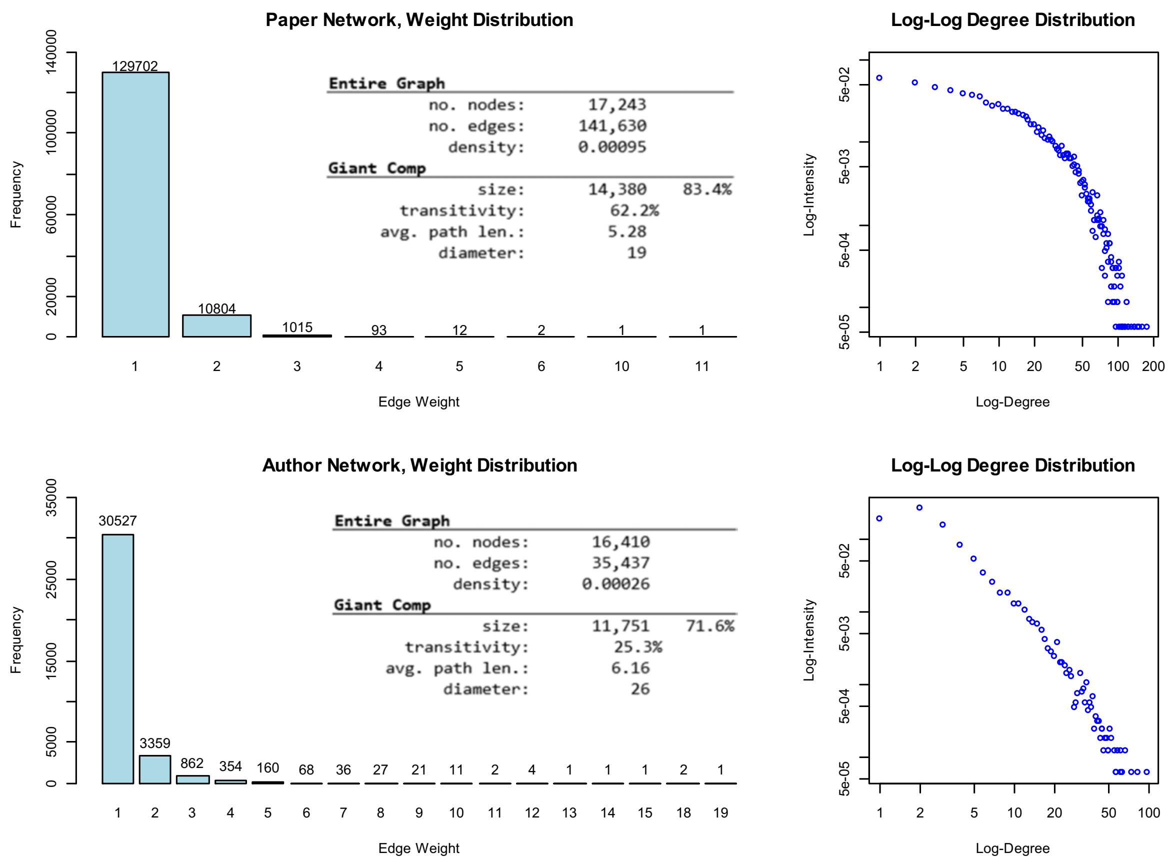

Fig. 8. Characteristics of the paper and author networks.

communities is sparse. A large portion of the intercommunity connections originates from a small number of papers. Because of its complexity, we present the annotated network diagram and list the pivotal papers in Web Appendix 4.

The Relationship Between Centrality and Citation

Inspired by Yan and Ding (2009), we also investigated the network attributes that may correlate to the number of citations. The results are summarized in Table 2, where citations (*ci*) and degree (*deg*), strange (*str*), betweenness (*bet*), and eigenvector centralities (*evc*) are log-transformed for normality. After the transformations, the median and mean of authors' citations are 3.178 and 3.197 respectively. Those of the papers are 2.565 and 2.578 respectively.

Models for authors' citations are listed on the left of Table 2. In the upper panel, we estimate each relevant factor's explanatory powers (R^2) and list them in descending order. The explanatory power of the network attributes (shaded in gray) is significant. Vertex strength (*str*: the number of papers with cooperative authors) and degree (*deg*: the number of unique coauthors) explain 36.3% and 30.4% of the variance in citations respectively, followed by betweenness (*bet*: the probability of the shortest paths between other authors (28.3%)). This result implies that "participating in collaborative research" and "cooperating with different partners" both increase citations, as does "bridging productive authors who do not cooperate directly".

When we apply the forward stepwise algorithm to these relevant factors (in the lower part of the table), the number of papers (*nop*) is selected first, followed by *bet* ($\Delta R^2 = .060$). The network attributes' joint explanatory power is 44.2% (model 16), which is higher than that of *nop* ($R^2 = .379$, model 9). Further, the former remains significant ($R^2 = .105$) even when the latter is controlled for (model 15). This finding indicates that good relationships in authors' social networks are positively related to higher citations.

On the right of Table 2, we investigate factors relevant to the papers' citations. The correlation between citations and network attributes is not as obvious in the papers' network as in the authors'. Collectively, network attributes only explain 3.89% of the variance (model 37), which decreases to 1.22% if year, journal, and topic (*tp*) are controlled for (model 36).

Discussion, Limitations, and Further Studies

In the spirit of a historical study, the way in which marketing has adapted to contemporary history concerns us. The Internet and globalization in the 1990s, and the burst bubble and financial crisis in the 2000s, brought dramatic changes to our business environment. The result of this study indicates that as the business environment and value system evolved, marketing adapted. During the past two decades, the focus of marketing has moved from product quality, customer satisfaction, and channel coordination to service dominance, brand differentiation, and customer

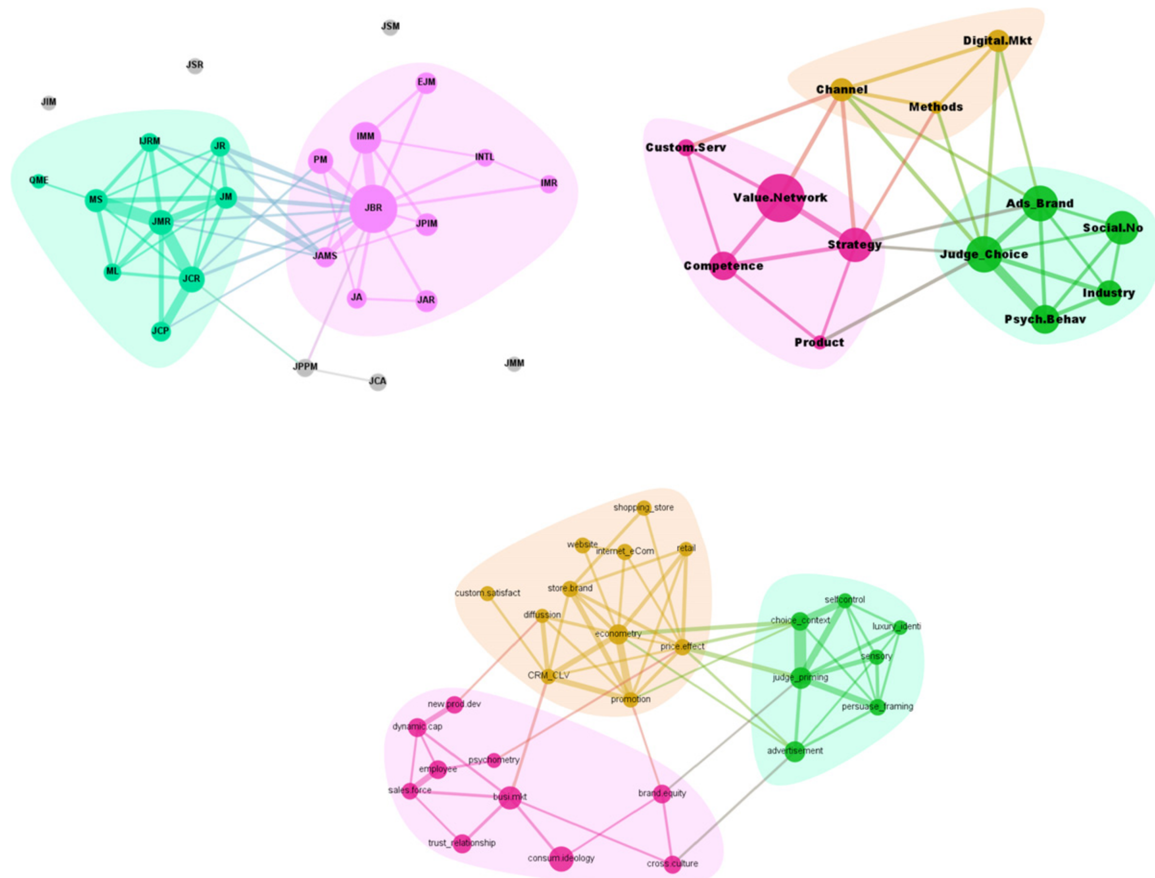

Fig. 9. a. Co-authorship network among journals. 1. Nodes are created by aggregating the papers by journals. 2. Node sizes and edge widths represent the number of papers in the journals and the number of co-authors between them. 3. For clarity, the edges of lower weight (<15% of the maximum) are censored. b. Co-authorship network among topic groups. 1. Node sizes and edge widths represent the number of papers in the topic groups and the number of co-authors between them. 2. For clarity, the edges of lower weights (<35%) are censored. c. Coauthorship network among topics. 1. Node sizes and edge widths represent the number of papers in the topics and the number of co-authors between them. 2. For clarity, the nodes of low eigenvector centralities (<50%) and the edges of low weights (<18%) are censored.

centrality. In the new millennium, while emerging markets have flourished in the midst of financial turmoil, the ethos of sustainability has been colliding into the once promising vision of a fully-digitized global village. As consumer conscience gradually overcomes materialism and consumerism, the practice of marketing has evolved from hedonic and individualistic appeals to the cocreation of experiential and social value.

Besides the aforementioned trends, this study also identifies several research areas undocumented in prior studies. “Value network” reflects the growing interest in network and platform-based business models. “Customer and service” not only identifies the varying implications of ‘customer’ (from CS, CRM, and CLV to CX), but also reveals its close relationship with ‘service’. The wax and wane of “cognitive behavior” and “behavioral psychologies” indicate that, in the long run, the purpose of marketing science is to serve consumers’ mental interests instead of exploiting their cognitive flaws.

We find that marketing research has not evolved in one direction. After moving toward strategic and cognitive topics, it returned to its traditional domains of customer service and consumer psychologies. Leveraging our universal corpus, we examine heterogeneity across 25 journals and visualize how topical composition and diversity may relate to impact factor and citations. Our findings show that the most impactful journals are the most diverse, whereas each runner-up has its own specialty. Besides the general topic groups, certain niche topic areas are captured by a few specialized journals who manage to acquire outstanding impacts.

Our social network analyses reveal that marketing researchers prefer to cooperate with new partners and that social capital in the authors’ network reinforces itself. Via community detection, we reveal the coauthorship structure among the journals, topic groups, topics, and the most cited papers. The pivots and bridging members in these communities are also identified. Indeed, this could be the first study that compares topical composition and analyzes the coauthorship structure across major marketing journals because such research is feasible only when topics can automatically be detected in a large scale.

Our study nonetheless has its limitations. Topic modeling demands domain knowledge as well as analytical skills. Regarding model selection, topic naming and topic grouping, because there is no standard and objective criterion to follow, we had to resort to our own professional judgment. Because other selecting, naming, and grouping schemes could lead to better results, we decided to share our topic model and its relevant data via GitHub for possible collaboration and further improvement.

Table 2
Models for the papers' and authors' citations.

Models the authors citations

no.	Models (remarks)	R^2	ΔR^2
1	*ci ~ nop* (number of papers)	37.85%	
2	*ci ~ str* (vertex stength)	36.31%	
3	*ci ~ deg* (degree)	30.36%	
4	*ci ~ bet* (betweenness)	28.29%	
5	*ci ~ wgt* (average weight)	15.24%	
6	*ci ~ clo* (closeness)	13.57%	
7	*ci ~ evc* (eigenvector centrality)	13.54%	
8	*ci ~* (null model)		
9	*ci ~ nop*	37.85%	37.85%
10	*ci ~ nop + bet*	43.83%	
11	*ci ~ nop + bet + wgt*	46.58%	
12	*ci ~ nop + bet + wgt + clo*	47.44%	
13	*ci ~ nop + bet + wgt + clo + str*	47.53%	
14	*ci ~ nop + bet + wgt + clo + str + deg*	48.34%	
15	*ci ~ nop + bet + wgt + clo + str + deg + evc*	48.38%	10.53%
16	*ci ~ str + deg + bet + clo + wgt + evc*	44.17%	

Models for the papers' citations

no.	Models (remarks)	R^2	ΔR^2
17	*ci ~ yr* (year)	31.61%	
18	*ci ~ jn* (journal)	13.65%	
19	*ci ~ tp* (topic)	3.87%	
20	*ci ~ deg* (degree)	3.57%	
21	*ci ~ str* (vertex strength)	3.57%	
22	*ci ~ clo* (closeness)	2.78%	
23	*ci ~ evc* (eigenvector centrality)	2.10%	
24	*ci ~ tg* (topic group)	1.14%	
25	*ci ~ bet* (betweenness)	0.98%	
26	*ci ~ noa* (number of authors)	0.25%	
27	*ci ~ wgt* (average weight)	0.01%	
28	*ci ~* (null model)	0	
29	*ci ~ yr*	31.61%	31.61%
30	*ci ~ yr + jn*	41.88%	10.27%
31	*ci ~ yr + jn + tp*	47.43%	5.55%
32	*ci ~ yr + jn + tp + str*	48.37%	
33	*ci ~ yr + jn + tp + str + deg*	48.54%	
34	*ci ~ yr + jn + tp + str + deg + evc*	48.64%	
35	*ci ~ yr + jn + tp + str + deg + evc + clo*	48.65%	
36	*ci ~ yr + jn + tp + str + deg + evc + clo + wgt*	48.66%	1.22%
37	*ci ~ deg + clo + bet + evc + str + wgt*	3.89%	

* Citations (*ci*) and the degree, strange, betweenness, and eigenvector centralities (*deg*, *str*, *bet*, and *evc*) are log-transferred for normality.

Recent advancements in data science are remarkable. In fact, some of the analytic tools adopted in this study were not available when Huber, Kamakura, and Mela (2014) conducted their research three years ago. Such extraordinary progress should be, at least partially, attributed to collaborative practice, whereby data scientists reproduce and improve each other's works on such online collaboration platforms as GitHub. We believe that, aside from analytical technologies, the practice of collaborative research could be the most valuable skill and mindset that we marketing researchers can learn from data scientists.

Appendix A. Topic Dictionary

Topics	Full name	Terms
ad.creativity	Ads Creativity	creativ/agenc/advertis agenc/client/advertis creativ/ad agenc/templat/creativ advertis/award/creativ process/creativ product/
ad.media	Ads Media	televis/media/commerci/televis advertis/audienc/viewer/program/broadcast/televis commerci/televis program/digit/radio/watch/s
advertisement	Advertisement	print/magazin/attitud advertis/print advertis/compar advertis/campaign/copi/advertis strategi/advertis product/aad/advertis c
age.groups	Age Groups	children/age/adolesc/older/parent/famili/adult/young/peer/teen/younger/older consum/teenag/girl/young adult/elder/peer influe
alliance	Strategic Alliance	allianc/subsidiari/partner/joint ventur/strateg allianc/ventur/foreign/ijv/fdi/brand allianc/intern joint/foreign direct/dire
arousal	Arousal	music/arous/pleasur/scent/ambient/nostalgia/pleasur arous/artist/art/color/background/ambient scent/nostalg/background music/
assortment_spatial [a]	Assortment & Spatial Models	assort/spatial/proxim/stockout/product assort/lotteri/assort size/ideal/attributebas/distanc/ideal point/space/geograph proxi
auction	Auction, Bidding	auction/bid/bidder/ship/onlin auction/reserv/seller/fee/ebay/bid behavior/revers auction/internet auction/herd/reserv price/s
brand.equity	Brand Equity	equiti/brand equiti/brand person/brand relationship/brand ident/brand loyalti/brand strategi/intern brand/consumerbrand/brand
brand.extens	Brand Extension	brand extens/parent brand/parent/cobrand/spillov/extens evalu/dilut/perceiv fit/fit brand/spillov effect/imag/extens product/
brand.image_COO [a]	Brand Image & Country of Origin	imag/coo/brand imag/ethnocentr/countri origin/countryoforigin/animos/consum ethnocentr/foreign/countri imag/domest/global bra

(*continued on next page*)

Appendix A (*continued*)

Topics	Full name	Terms
busi.mkt	B2B & Industrial Markets	busi market/evolutionari/industri market/agenda/improvis/market plan/businesstobusi market/ict/theori market/realism/market p
busi.network	Business (Rel.) Network	network/actor/busi network/social network/imp/network pictur/busi relationship/relationship network/busi model/embedded/indus
buy.sell.relation	Buyer-Seller Relationship	buyer/seller/buyersel/buyer seller/buyersel relationship/negoti/buy center/industri buyer/emarketplac/organiz buy/buyer behav
celebr.endors	Celebrity Endorsement	endors/celebr/credibl/celebr endors/brand credibl/colleg/advocaci/colleg student/matchup/endors brand/corpor credibl/endors e
channel	Channel Management	channel/distributor/market channel/channel relationship/channel member/multichannel/channel distribut/resel/intermediari/infl
china	China	china/chines/institut/guanxi/transit/transit economi/polit/russian/institut theori/institut environ/republ/chines firm/vietna
choice_context	Choice & Context Effect	option/choic set/compromis/regret/attract effect/effect choic/compromis effect/choic process/context effect/nonalign/choic co
claim_humor [a]	Ads Claims & Humor	claim/humor/skeptic/violenc/advertis claim/lowinvolv/ad claim/elm/violent/peripher/humor advertis/likelihood model/mislead/pa
collabor_KM [a]	Collaboration & Knowledge Management	collabor/transfer/cooper/knowledg transfer/depart/tacit/coopetit/knowledg manag/tacit knowledg/knowledg creation/knowledg sha
community_mobile	Mobile & Online Communities	communiti/mobil/user/brand communiti/virtual/lead user/onlin communiti/mobil phone/tam/toolkit/mass custom/perceiv use/sms/pe
congruency	(Brand) Congruency	incongru/languag/name/brand name/congruiti/congruent/implicit/bilingu/english/schema/linguist/accent/sound/moder incongru/con
conjoint.anal	Conjoint Analysis	conjoint/conjoint analysi/concept test/choicebas/rule/partworth/attribut level/choicebas conjoint/screen/weight/choic experi/
construal_focus [a]	Self Construal & Regulatory Focus	regulatori/construal/regulatori focus/tempor/construal level/preventionfocus/concret/abstract/frame/promotionfocus/regulatori
consum.ideology	Customer Ideology	consumpt/subsist/ideolog/authent/ethnograph/discours/introspect/ident/poverti/anticonsumpt/ritual/stori/narrat/resist/myth/so
credit.card_debt	Credit Card & Debt	credit/card/credit card/debt/loan/retir/bank/selfservic/sst/payment/financi literaci/literaci/save/selfservic technolog/finan
crisis_IMC	Crisis & Integrated Marcom.	crisi/market communic/imc/recess/audit/kingdom/unit kingdom/integr market/cyclic/nich/crise/public relat/disast/communic imc/
CRM_CLV	CRM & CLV (Custom. Life-time Value)	crm/retent/lifetim/clv/custom retent/custom lifetim/custom profit/defect/churn/crosssel/acquisit/custom acquisit/custom infor
cross.culture	Cross Culture	crosscultur/japanes/nation cultur/crossnat/japan/american/cultur differ/cultur valu/collectivist/korean/asian/hofsted/individ
CSR_reputation	CSR & Reputation	corpor/csr/corpor brand/social respons/corpor social/reput/corpor imag/corpor reput/corpor ident/csr activ/irrespons/corpor c
custom.satisfact	Customer Satisfaction	satisfact/custom satisfact/repurchas/consum satisfact/servic qualiti/satisfact loyalti/disconfirm/delight/repurchas intent/ef
diffusion	Diffusion Theory/Model	diffus/preannounc/takeoff/innov/diffus model/softwar/neural network/network extern/neural/agentbas/bass/diffus process/hardwa
donation	Donation, Charities	donat/cycl/cycl time/chariti/charit/donor/direct mail/mail/volunt/life cycl/causerel/causerel market/prosoci/develop cycl/mor
dynamic.cap	Dynamic Capability	capabl/turbul/market capabl/orient/innov/dynam capabl/innov capabl/organiz learn/ambidexter/technolog capabl/strateg orient/e
econometry	Econometrics	household/error/state depend/algorithm/stochast/market structur/probit/price sensit/parametr/respons model/threshold/probit m
emotion	Emotions	emot/negat emot/cope/boycott/posit emot/apprais/emot respons/consum emot/pride/quo/status quo/regret/emot appeal/attach/emot
employee	Employee & Job Satisfaction	employe/job/job satisfact/frontlin/servic employe/citizenship/frontlin employe/citizenship behavior/climat/satisfact/organiz
environment	Environment & Sustainability	environment/green/sustain/recycl/ecolog/super/bowl/green product/environment concern/green market/super bowl/energi/green adv
ethnics_minors	Ethnic Groups & Minorities	ethnic/minor/black/hispan/gay/accultur/white/religi/immigr/african/religios/africanamerican/african american/racial/multicult
export_psycho.dist [a]	Exportation & Psychological Distance	export/export market/export perform/psychic/export ventur/psychic distanc/distanc/firm export/export sale/export inform/emo/e
fail.recovery	Failure Recovery	recoveri/failur/justic/servic failur/complaint/servic recoveri/fair/complain/failur recoveri/distribut justic/dissatisfact/pr
food	Foods	food/nutrit/label/health/health claim/nutrit inform/obes/calori/nutrit label/fat/meal/nutrient/food product/claim/fast food/r
franchise_TCE [a]	Franchise & Transaction Cost Eco.	franchis/franchise/transact cost/franchisor/contract/licens/asset/franchis system/cost econom/contractu/tce/monitor/cost anal

Appendix A (*continued*)

Topics	Full name	Terms
gender.diff	Gender Difference	gender/women/femal/gift/male/men/sexual/sex/men women/gender differ/male femal/feminin/gift give/masculin/portray/giver/giftg
health.care	Health Care	health/drug/physician/prescript/health care/patient/medic/pharmaceut/healthcar/dtc/directtoconsum/prescript drug/dtca/hospit/
hedonic_utilitar	Hedonic, Utilitarism	hedon/utilitarian/experienti/hedon utilitarian/hedon product/utilitarian product/consumpt/utilitarian hedon/selfexpress/psych
hongkong_NPO [a]	Hong Kong & Non-Profit Organization	hong/kong/hong kong/nonprofit/trade show/trade/nonprofit organ/museum/organis/forprofit/handbil/intimaci/standardis/npo/cogni
innovation	Innovation	innov/radic/product innov/radic innov/innov perform/servic innov/patent/firm innov/organiz innov/open innov/innov success/bre
internet_eCom	Internet & E-Com.	onlin/internet/offlin/onlin shop/onlin retail/shop/ecommerc/retail/etail/onlin store/multichannel retail/internet retail/onli
intl_SMES	International SMB/SME's	internation/smes/mode/entrepreneuri/internationalis/foreign/entri mode/entri/foreign market/small firm/domest/entrepreneurshi
judge_priming	Judgment and Priming	judgment/mood/prime/assimil/consum judgment/anchor/posit mood/product evalu/automat/mindset/heurist/contrast effect/spontan/c
key.acc.mgn	Key Account Management	key account/account manag/kam/tactic/intellig/empower/align/market intellig/organis/competit intellig/marketingsal/influenc t
leadership	Leadership	leadership/ceo/board/transform leadership/transform/style/tmt/director/chief/manag team/leadership style/chief execut/tenur/c
loss.averse	Loss/Risk Aversion	wait/avers/loss/contest/incent/loss avers/risk avers/wait time/endow/endow effect/risk/winner/compens/prize/sale contest/quot
loyalty	Loyalty	loyalti/switch/custom loyalti/switch cost/commit/attitudin/affect commit/brand loyalti/attitudin loyalti/switch behavior/bank
luxury_identity [a]	Luxury & Social Identity	luxuri/ident/luxuri brand/selfconstru/social ident/counterfeit/status/identif/fashion/conspicu/consumpt/outgroup/interdepend/
macro.mkt_busi.edu [a]	Macro-Marketing & Business Education	macromarket/busi school/institut/societi/professor/special section/cours/book/lead market/market educ/market associ/metric/su
mail_survey	(Mail) Survey	mail/mail survey/norm/respons rate/selfcongru/plan behavior/theori plan/subject norm/selfimag/social norm/followup/perceiv co
memory	Memory	memori/recal/exposur/placement/recognit/familiar/game/retriev/product placement/interfer/rememb/brand recal/repetit/consum me
meta.analysis	Meta Analysis	metaanalysi/ventur/empir general/effect size/escal/metaanalyt/commit/inform acquisit/market inform/escal commit/ra theori/the
mkt.entry	Market Entrance	entri/pioneer/incumb/entrant/defens/surviv/disrupt/market entri/exit/pioneer advantag/firstmov/mover/disrupt innov/order entr
movies	Movies	movi/forecast/film/releas/motion/motion pictur/box/box offic/pictur/sale forecast/offic/sequel/theater/studio/pictur industri
new.prod.dev	New Product Development	npd/project/launch/innov/profici/speed/product perform/product innov/champion/radic/new/product advantag/front end/innov proj
persuase_framing	Persuasion & Framing	messag/persuas/frame/persuas knowledg/narrat/ambival/elabor/messag frame/nfc/selfreferenc/appeal/favor attitud/cognit respons
price.effect	(Psych.) Price Effect	bundl/refer price/discount/price strategi/price increas/price inform/partit/price effect/price percept/unfair/intern refer/su
privacy_fair.trade [a]	Privacy & Fair Trade	privaci/latin/america/latin america/privaci concern/fair trade/latin american/certif/costa/trade/person inform/consum privaci
prod.knowledge	Product Familiarity/Expertise	expert/cue/novic/trial/analog/familiar/categor/product trial/extrins/expertis/product knowledg/knowledg consum/halo/consum kn
promotion	(Price) Promotions	coupon/elast/sale promot/gambl/price promot/price elast/discount/free/redempt/crosspric/prone/promot strategi/crossbrand/casi
psychometry	Psychometry	scale/scale develop/psychometr/reliabl valid/multidimension scale/develop scale/measur construct/psychometr properti/respons
regulation	Regulation	legal/regul/law/protect/disclosur/feder/decept/commiss/trademark/vulner/court/feder trade/trade commiss/antitrust/legisl/disc
retail	Retail (vs. Wholesale)	retail/guarante/slot/wholesal/passthrough/pricematch/wholesal price/return polici/trade promot/channel/refund/trade/pricematc
rewarding	Rewarding, Loyalty Programs	program/reward/loyalti program/loyalti/market program/reward program/referr/membership/product program/program effect/incent/
risky_habit [a]	Bad Habit & Risk Perception	risk/smoke/perceiv risk/alcohol/warn/cigarett/tobacco/smoker/risk percept/social market/drink/antismok/riski/campaign/risktak
sales.force	Sales Force	salesperson/salespeopl/sale manag/sale forc/sale perform/train/salesperson perform/sfa/sale organ/adapt sell/salesforc/sale t
SCM_logistics	Supply Chain Mgn. & Logistics	suppli/chain/suppli chain/logist/chain manag/logist servic/lean/scm/ecr/revers logist/suppli side/manag suppli/global suppli/

(*continued on next page*)

Appendix A (*continued*)

Topics	Full name	Terms
search.ads	Search Engine/Ads	search engin/search/keyword/quantiti/engin/click/crosscategori/purchas incid/search advertis/purchas quantiti/incid/shelf/gro
search_automobile [a]	Online Inform. Search for Automobiles	search/inform search/car/automobil/search cost/consum search/search behavior/credenc/dealer/leas/price search/recommend agent
segmentation	Segment, Cluster, Classification	segment/cluster/classif/criteria/mixtur/psychograph/finit mixtur/taxonomi/lifestyl/segment model/fuzzi/scheme/select criteria
selfcontrol	Self-control	consumpt/impuls/selfcontrol/food/indulg/deplet/impuls buy/eat/packag/temptat/healthi/satiat/virtu/selfregul/vice/increas cons
sensory	Sensory & Aesthetics	visual/aesthet/touch/metaphor/imageri/logo/imagin/color/verbal/pictori/haptic/rhetor/represent/graphic/embodi/packag/product
service_cocreat	Service Logic & Co-Creation	servic qualiti/cocreat/logic/servicedomin/nsd/servic develop/eservic/servicedomin logic/sd logic/servic product/servicescap/c
shopping_store	Shopping & Storefronts	shop/store/shopper/mall/retail/patronag/instor/trip/shop trip/shop behavior/basket/shop experi/groceri/store choic/shop motiv
signal_family [a]	Signaling & Family Businesses	signal/diversif/famili/busi group/famili firm/sign/sale sign/famili busi/signal qualiti/highqual/signal product/encroach/qual
soc.capit_prod.dgn [a]	Social Capital & Product Design	capit/social capit/solut/engin/modular/platform/architectur/product design/product famili/problemsolv/integr solut/reus/produ
soc.media_politics [a]	Social Media & Politics	polit/social media/voter/elect/viral/opinion/media/blog/usergener/cynic/vote/ugc/candid/campaign/opinion leader/usergener con
spend.money_time	Spending Time/Money	spend/money/budget/pet/spent/dog/dollar/currenc/mental/consum spend/mental account/spend time/time money/denomin/save money/s
sponsorship	Sponsorship	sponsorship/sponsor/event/sport/fan/ambush/sport event/ambush market/sponsor event/identif/sponsorship effect/olymp/sponsor b
stakehold_wellbeing	Stakeholder Wellbeing	stakehold/wellb/intervent/modif/multipl stakehold/market system/dark side/stakehold market/societ/welfar/hous/stakehold theor
stock.return	Stock and ROI	stock/investor/sharehold/announc/stock market/fund/cash/stock return/acquisit/merger/earn/cash flow/stock price/abnorm/asset/
store.brand	(Retail) Store Brand	retail/store/store brand/privat label/nation brand/supermarket/privat/groceri/walmart/label/merchandis/intern retail/retail f
strategic.orient	Market & Customer Orientation	orient/market orient/ethic/custom orient/marketori/competitor orient/uneth/orient product/strateg orient/entrepreneuri orient
supply_outsource	Supplier & Outsourcing	supplier/outsourc/buyersuppli/buyersuppli relationship/supplier relationship/supplier involv/supplier perform/buyer/key suppl
teamwork	(Cross-Function) Teamwork	team/crossfunct/conflict/project/npd/npd team/crossfunct integr/team perform/sale team/dispers/rdmarket/conflict manag/interf
trait_personality	Personal Trait & Personality Appeal	trait/appeal/person trait/scarciti/fear/selfmonitor/person valu/maven/materi/superstit/fear appeal/market maven/personif/guil
travel_tourism	Travel, Tourism	hotel/tourism/travel/destin/tourist/map/airlin/meansend/ladder/gray/guest/gray market/envelop/cognit map/data envelop/envelop
trust_relationship	Trust & Relationship	trust/commit/trust commit/relationship qualiti/norm/interperson/exchang relationship/relat exchang/interorganiz/consum trust/
two.part.tariff	(Two-Part) Tariff	usag/tariff/attain/product usag/twopart/smart/twopart tariff/pursuit/heavi user/user/goal pursuit/consum goal/usag situat/hea
website	Websites	web/site/internet/web site/websit/banner/onlin/email/page/banner advertis/brows/wide web/internet user/world wide/user/banner
WoM	Word of Mouth	wom/mouth/word mouth/tie/referr/negat word/nwom/pwom/posit word/social comparison/negat wom/wordofmouth communic/posit wom/ti
WTP_price	WTP & Price Discrimination	wtp/willing pay/price discrimin/valuat/durabl/product line/premium/resal/warranti/monopolist/durabl good/ration/upgrad/welfar

Compound topics	Full name	Connection
assortment_spatial	Assortment & Spatial Models	Spatial models are widely used for product portfolio and shelf assortment.
brand.image_COO	Brand Image & Country of Origin	Country of Origin is a kind of brand.
claim_humor	Ads Claims & Humor	These papers discuss the effects of claims and humors embedded in ads.
collabor_KM	Collaboration & Knowledge Management	Collaboration (and cross functional cooperation) is crucial for Knowledge Management (KM).
construal_focus	Self Construal & Regulatory Focus	Self-perception (self-construal) influences the orientation of self-regulation.
export_psycho.dist	Export & Psychological Distance	Psychological/cultural distance often concerns exporters.
franchise_TCE	Franchise & Transaction Cost Eco.	Many Franchise systems are designed on the framework of Transaction Cost Economics.
hongkong_NPO	Hong Kong & Non-Profit Organization	Many NGO concerning child labor exploitation in Asia are hosted in Hong Kong.
luxury_identity	Luxury & Social Identity	Luxury goods can signal Social Status.
macro.mkt_busi.edu	Macro-Marketing & Business Education	The ethics of macro-marketing should be cultivated in business schools.
privacy_fair.trade	Privacy & Fair Trade	Both privacy and fair trade concern business ethics.

Appendix A (*continued*)

Compound topics	Full name	Connection
risky_habit	Bad Habit & Risk Perception	Risk perception might suppresses habitual consumption of alcohol, tobacco, etc.
search_automobile	Inform. Search & Auto Industries	Car buyers often search for information online.
signal_family	Signaling & Family Businesses	Cannot identify plausible connection [b]
soc.capit_prod.dgn	Social Capital & Product Design	Cannot identify plausible connection [b]
soc.media_politics	Social Media & Politics	Social media is heavily used in political marketing.

[a] Connections between the elements of compound topics are explained below.
[b] Cannot identify plausible connections, it seems that the elements are simply close to each other in the posterior probability space.

Table A5.1
Modeling the trend of topic groups.

	Model.5	Model.4	Model.3	Model.2	Model.1	Model.0
	Adj. AR(1) & Heteroscedasticity	3rd order Random slope	2nd order Random slope	1st order Random slope	Random intercept	Null model
Model fitness (*p*-value [a])						
LogLik	841 (1e−4)	834 (0.70)	833 (<0.001)	816 (<0.0001)	762 (1)	762
AIC	−1658 (1e−4)	−1639 (0.70)	−1646 (<0.001)	−1619 (<0.0001)	−1515 (1)	−1517
Fixed effect, beta (*p*-value)						
Intercept	0.0769 (0)	0.0769 (0)	0.0769 (0)	0.0769 (0)	0.0769 (0)	0.0769 (0)
Random effect, std.deviation						
Intercept	0.01657	0.01655	0.01657	0.02250	0.01645	0.01645
Time	0.12399	0.12380	0.12388	0.00132		
Time^2	0.06033	0.06270	0.06263			
Time^3		0.01409				
Residual, %total.var.	27.34%	14.62%	14.83%	18.90%	33.81%	33.81%
Auto-correlation						
Phi (AR1)	0.023					

[a] The *p*-value is estimated by ANOVA against the previous model.

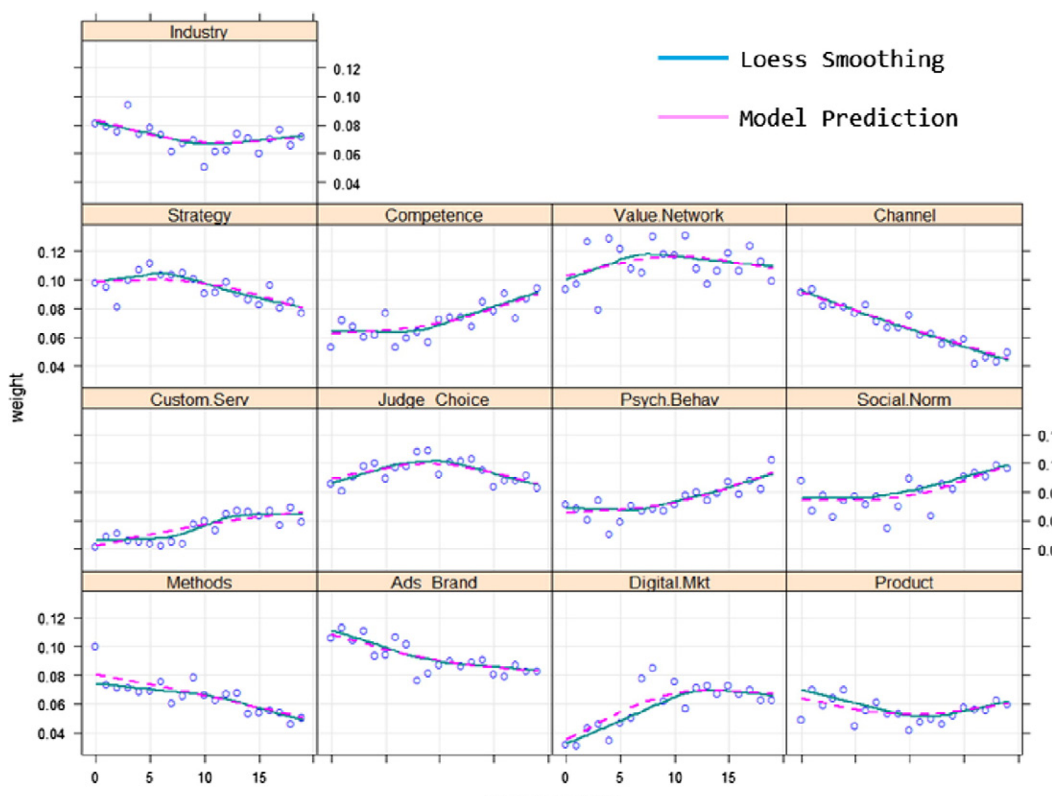

Fig. A5.1. Comparing model prediction with loess smoothing.

Appendix B. Hierarchical Regression Model for the Trend of Topic Groups

To examine the trend of the topics groups, we analyze the groups' annual weight with hierarchical regression models. First, we aggregate the weights of papers by topic groups and normalize the sum of annual weights to 1. Thereby, we have a time series of 20 periods for each topic group. We then follow Bliese and Ployhart (2002) to build a random effect model for hierarchical time series data. In the following process, we basically aim to explain the variance of 'weight' by the fixed and random effects of 'time' and its higher order forms.

B.1. Null Model (model.0) and Random Intercept (model.1)

First, we examine the interclass correlation (ICC) with a null model that only has a random intercept for each topic group. It shows that the random intercept explains 66% of the total variance, which is more than enough to justify random effect model.

B.2. 1st Order Random Slope (model.2)

Model fitness significantly improved ($p<0.0001$, according to ANOVA) as we adopt the 1st order random slope. However, the random slope does not explain much variance (only 0.3%) by itself. Because we have normalized the time series by period, there is no fixed effect on time either.

B.3. 2nd (model.3) and 3rd (model.4) Order Random Slope

The 2nd order random slope significantly improves model fitness ($p<0.0001$, model.2 against model.3). The 1st and 2nd order random slopes jointly explain 81.1% of the variance. But, adding the 3rd order random slope does not improve the model ($p=0.701$, model.3 against model.4).

B.4. Adjust for Auto-correlation and Heteroscedasticity (model.5)

Therefore, we fallback to model.3 and then adjust our model for the time series specific dependencies. It appears that auto-correlation is low (AR(1) = 0.023), but there is a trend of decreasing variance over time. Adjusting for these dependencies does improve the fitness significantly ($p<1\ e^{-4}$). The variance of residual rose to 27.34% after the adjustment. It indicates that standard ML estimator would over-estimate the accuracy; adjustment is thus required.

The above modeling process is summarized in Table A5.1. In Fig. A5.1, we compare the prediction of the final model (pink dash lines) with the Loess smoothing curves (light-blue solid lines) automatically generated by the plotting tool. It seems that the former is no better than the latter. In such topic groups as "customer and service," moving average fits the data better than hierarchical regression does.

Appendix C. LDA Parameters Tuning Process

There are three major parameters in LDA-based topic models. K determines the number of topics. α and η control the initial distribution of topics and terms within documents and topics

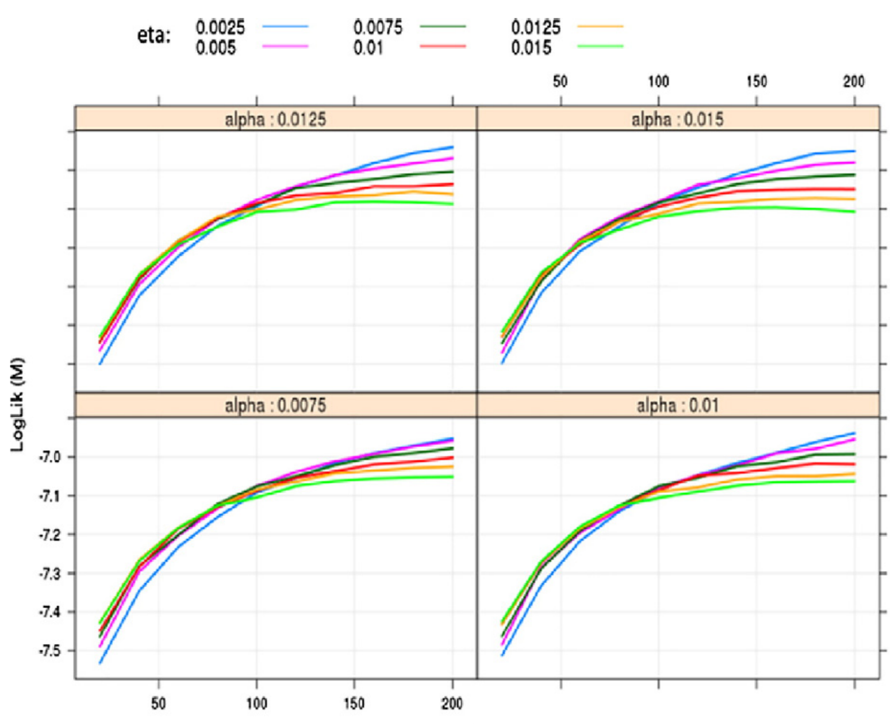

Fig. A6.1. Extended grid exploration.

respectively. Smaller αs and ηs lead to flatter distributions. Other than that α should approximately be the reverse of K, there is no standard rule for determining the parameters. We adopt a four-stage parameter tuning and model selection process as explained below.

C.1. Exploration

We start by a broad exploration in the parameter space $0.002 \leq \alpha$, $\eta \leq 0.2$ and $20 \leq K \leq 300$. Based on the log-likelihood and the convergence situation of the models, we narrow down the parameter space and perform an extended-grid scanning. The result is shown in Fig. A6.1. Disregarding α and η, it appears that the log-likelihood starts fattening when $K \geq 80$.

C.2. Candidate Models

Like all of the other machine learning technologies, the LDA models that exhibit higher likelihood may not always make better sense to human (Chang et al. 2009). Knowing that the quality of topic modeling cannot solely be determined by the fitness statistics (Sievert and Shirley 2014), we select a set of 24 candidate models, where $K \in \{60, 70, 80, 90, 100, 110, 120\}$, $\alpha = 1/K$, and $\eta \in \{0.005, 0.0075, 0.01, 0.015\}$, based on the result of the exploration stage.

C.3. Manual Screening

Each of the 24 candidates defines each of its K topics with a list of terms and each of the topics has to be manually named. As K increases, we find more topics with more precise definitions. Some topics may look ambiguous at the first glance. However, when we look closer, we find that the seemingly unrelated elements in most of these 'compound' topics, such as 'search & automobile' and 'export & distance,' do have a good reason to combine. (Please refer to the list of compounded topics at the bottom of Appendix A.) Since our objective is to extract as many clearly defined topics as possible, we choose to stop at $K = 100$, where ambiguous topics, such as 'signal & family business' and 'social capital & product design,' start to appear.

C.4. Final Selection

Fixing at $K = 100$, we do a final scan for the best combination of α and η. While subjective judgment cannot be excluded, we make our final selection by the following criteria:

1. All topics should be salient and distinctive.
2. A topic should cover either a specific subject or a set of related subjects.
3. A tradeoff between comprehension and distinctiveness is performed at the model level.

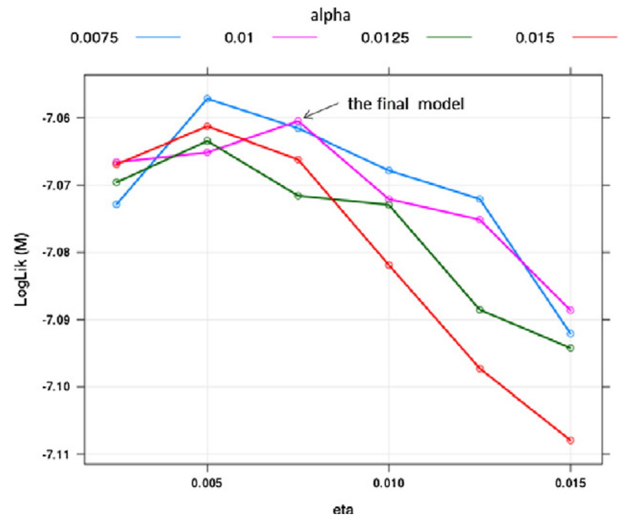

Fig. A6.2. Extended grid exploration.

As shown in Fig. A6.2, the final model is at (α, η) = (0.01, 0.0075). Although not the most likely, it captures many salient and distinctive topics, such as 'sensory (marketing)' and 'word of mouth,' and only contains two ambiguous topics, 'signal & family (business)' and 'social capital & product design'.

Appendix D. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.intmar.2017.06.003.

References

Ariely, Dan (2008), *Predictably Irrational*. New York: HarperCollins.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003, Jan.), "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Bliese, Paul D. and Robert E. Ployhart (2002), "Growth Modeling Using Random Coefficient Models: Model Building, Testing, and Illustrations," *Organizational Research Methods*, 5, 4, 362–87.

Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre (2008), "Fast Unfolding of Communities in Large Networks," *Journal of Statistical Mechanics: Theory and Experiment*, 2008, 10P10008.

Chang, Jonathan, Jordan L. Boyd-Graber, Sean Gerrish, Chong Wang, and David M. Blei (2009), "Reading Tea Leaves: How Humans Interpret Topic Models," *Advances in Neural Information Processing Systems*, 31, 1–9.

——— (2015), "lda: Collapsed Gibbs Sampling Methods for Topic Models," *R Package, Version 1.4.2.* (http://cran.r-project.org/web/packages/lda/lda.pdf).

Chowdhury, Gobinda (2010), *Introduction to Modern Information Retrieval*. Facet Publishing.

Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman (1990), "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, 41, 6, 391–407.

Goldenberg, Jacob, Barak Libai, Eitan Muller, and Stefan Stremersch (2010), "Database Submission — The Evolving Social Network of Marketing Scholars," *Marketing Science*, 29, 3, 561–7.

Gower, John C. (1966), "Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis," *Biometrika*, 53, 3/4, 325–38.

Hornik, Kurt and Bettina Grün (2011), "Topicmodels: An R Package for Fitting Topic Models," *Journal of Statistical Software*, 40, 13, 1–30.

Huber, Joel, Wagner Kamakura, and Carl F. Mela (2014), "A Topical History of JMR," *Journal of Marketing Research*, 51, 1, 84–91.

Lemon, Katherine N. and Peter C. Verhoef (2016), "Understanding Customer Experience Throughout the Customer Journey," *Journal of Marketing*, 80, 6, 69–96.

Leydesdorff, Loet and Ismael Rafols (2011), "Indicators of the Interdisciplinarity of Journals: Diversity, Centrality, and Citations," *Journal of Informetrics*, 5, 1, 87–100.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze (2008), *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

Mela, Carl F., Jason Roos, and Yiting Deng (2013), "Invited Paper — A Keyword History of Marketing Science," *Marketing Science*, 32, 1, 8–18.

Meyer, David, Kurt Hornik, and Ingo Feinerer (2008), "Text Mining Infrastructure in R," *Journal of Statistical Software*, 25, 5, 1–54.

Newman, Mark E.J. (2001), "The Structure of Scientific Collaboration Networks," *Proceedings of the National Academy of Sciences*, 98, 2, 404–9.

Porter, Martin F. (1980), "An Algorithm for Suffix Stripping," *Program*, 14, 3, 130–7.

Sievert, Carson and Kenneth E. Shirley (2014, June), "LDAvis: A Method for Visualizing and Interpreting Topics," *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. 63–70.

Vargo, Stephen L. and Robert F. Lusch (2008), "Service-dominant Logic: Continuing the Evolution," *Journal of the Academy of Marketing Science*, 36, 1, 1–10.

Wang, Xin Shane, Neil T. Bendle, Feng Mai, and June Cotte (2015), "The Journal of Consumer Research at 40: A Historical Analysis," *Journal of Consumer Research*, 42, 1, 5–18.

Wild, Fridolin, Depends Rstem, and Maintainer Fridolin Wild (2009), "An LSA Package for R," *Mini-proceedings of the 1st European Workshop on Latent Semantic Analysis in Technology-enhanced Learning*. 12.

Wong, Wilson, Wei Liu, and Mohammed Bennamoun (2012), "Ontology Learning from Text: A Look Back and into the Future," *ACM Computing Surveys (CSUR)*, 44, 4, 20, 1–36.

Yan, Erjia and Ying Ding (2009), "Applying Centrality Measures to Impact Analysis: A Coauthorship Network Analysis," *Journal of the Association for Information Science and Technology*, 60, 10, 2107–18.