J. of the Acad. Mark. Sci. (2015) 43:52–72
DOI 10.1007/s11747-014-0392-7

ORIGINAL EMPIRICAL RESEARCH

Explaining article influence: capturing article citability and its dynamic effects

Shibo Li · Eugene Sivadas · Mark S. Johnson

Received: 10 October 2013 / Accepted: 25 April 2014 / Published online: 28 May 2014

Abstract Researchers from diverse disciplines have examined the many factors that contribute to the influence of published research papers. Such influence dynamics are in essence a marketing of science issue. In this paper, we propose that in addition to known established, overt drivers of influence such as journal, article, author, and Matthew effects, a latent factor "citability" influences the eventual impact of a paper. Citability is a mid-range latent variable that captures the changing relationship of an article to a field. Our analysis using a discretized Tobit model with hidden Markov processes suggests that there are two states of citability, and these dynamic states determine eventual influence of a paper. Prior research in marketing has relied on models where the various effects such as author and journal effects are deemed static. Unlike ours, these models fail to capture the continuously evolving impact dynamics of a paper and the differential effect of the various drivers that depend on the latent state a paper is in at any given point of time. Our model also captures the impact of uncitedness, which other models fail to do. Our model is estimated using articles published in seven leading marketing journals during the years 1996–2003. Findings and implications are discussed.

Keywords Citation · Citability · Matthew effect · Uncitedness · Scientometrics · Hidden Markov model · Tobit

Introduction

The perceived quality of research is important both for individual scholars and for journals as they are intrinsically linked to the standing of both and are at essence a "marketing of science" issue (Stremersch et al. 2007; Varadarajan 2003). Two widely used approaches to assess quality are to look at either an input measure (the perceived quality or ranking of the journal the article was published in) or output measures (quality assessment via the influence and impact of an article's accrued citations) (Bergh et al. 2006; Garfield 1979; Medoff 2006). Both these approaches exemplify the "peer recognition of an academic's research" which is central to the academic reward system (Medoff 2006).

However, looking at journal reputation, rankings, or journal impact factor is problematic as a measure of individual article influence as many articles published in influential journals may not be influential, and influential articles are also published in relatively less prestigious journals. As Seglen (1997) and Woodside (2009) argue, journal impact factor metrics are unrepresentative of most of the articles published in academic journals and a "poor proxy" for the actual impact a specific article might generate. The San Francisco Declaration on Research Assessment (DORA) initiated by the American Society of Cell Biology in 2012 also argues that the Journal Impact Factor is a poor indicator of scientific quality of an individual article (http://am.ascb.org/dora/). The San Francisco declaration further urges readers to "assess research on its own merit."

In this paper, we propose that to completely capture and understand the reasons for an article's influence, one needs to assess *article citability*. We develop a method to gauge this

Mark S. Johnson is deceased

S. Li
Kelley School of Business, Indiana University, 1309 East Tenth Street, Bloomington, IN 47405, USA
e-mail: shili@indiana.edu

E. Sivadas (✉)
Milgard School of Business, University of Washington, Tacoma, 1900 Commerce St, Campus Box 358420, Tacoma, WA 98402, USA
e-mail: sivadas@u.washington.edu

latent construct. We suggest that continually evolving citability explains article influence measured using citation counts, and we shed light on the dynamic change of article citability on impact over time after publication. We capture the moderating effect of aggregate article citability on various known drivers of impact and suggest that citability of a paper dynamically changes over time; this citability determines overall article influence.

Our research builds upon extant work and contributes to the understanding of marketing scholarship in specific and academic scholarship in general in the following ways. We introduce to the literature a construct we term "article citability." The citability construct captures a discipline's and aggregate scholars' level of interest in a paper. We treat this interest as dynamic and ever changing and therefore develop a dynamic model that can capture this latent construct. We treat article citability as a mid-range latent variable that taps into the relational embededness of a paper to a discipline, as well as the interest a paper might generate within a discipline or sub-discipline due to its brand equity or quality. Notions of quality and brand equity are therefore subsumed in the citability construct. To capture article level dynamics we adopt a modified version of the count model—a discretized Tobit model with a hidden Markov process proposed by Li et al. (2005). Our model incorporates key influencers of influence and controls for time and unobserved article heterogeneity not encapsulated by previous research (Burrell 2003; Stremersch et al. 2007).

We posit that an article's influence dynamics are driven by the latent citability of the work and the discipline's relationship with a paper, which dynamically changes over time. By capturing these dynamics, we demonstrate that various effects that drive impact (such as author, article, and journal) are not static but change over time. We demonstrate that our model outperforms the commonly used NBD model, a static model which assumes that the various effects are static and do not change over time. We demonstrate that various influencers of impact change according to the latent citability states, and citability states of an article also changes dynamically over time.

There has been some recognition amongst scientometric scholars that the works of scholars and articles with greater visibility are more likely to be noticed and cited. This phenomenon is termed the Matthew effect (Merton 1968). However, prior research in marketing (e.g., Stremersch et al. 2007) has conceptualized the Matthew effect narrowly, simply accounting for the halo of author prestige and its effect on citations based impact. We expand on the conceptualization of the Matthew effect and suggest that the Matthew effect is not simply a famous author effect but a much larger brand equity signaling a famous paper effect mechanism. The visibility of the work itself or what may be characterized as the famous paper effect is captured in our model. We find that famous articles have a significant Matthew effect independent of the famous author effect that generates a halo and results in citations that go above and beyond what can be explained by the quality of the work itself.

We provide new and deeper insights into assessing article citability and capturing the dynamic nature of such an enterprise. Additionally, we capture the effect of uncitedness that most models fail to capture. We next introduce our conceptual framework. We then discuss our data and model, and provide empirical results and outline implications and theoretical contributions. Finally, we conclude with research limitations and future directions.

Conceptual framework

Not surprisingly, the topic of quality and influence of published research—scientometrics and bibliometrics—has attracted the attention of a large and diverse group of scholars from a variety of disciplines: social, information, life, and engineering sciences. Journal publishers and editors, librarians in charge of making collections decisions, and individual scholars as well as administrators and others evaluating them for hiring, tenure, and promotion are interested in issues surrounding the quality and impact of published research. Thus, it is not surprising that research on this topic, while widespread and multidisciplinary, is rather fragmented.

In the marketing area, Stremersch et al. (2007) suggest that the influence an article generates can be explained by three differing theoretical viewpoints. The first, which they term the "universalist perspective," suggests that an article is cited because of its content and what it says. The universalist perspective comprises two dimensions: article quality and article domain. The second perspective is termed "the social constructivist perspective" and suggests that articles are cited because of the fame of the authors and personal promotion of articles by them. The third perspective, "presentation perspective," indicates that an article is influential for "how" the authors say what they say (p. 174). The presentation perspective suggests that the article title, keywords, and expositional clarity would influence citations. They found support for the universalist perspective and partial support for the social constructivist perspective but very little support for the presentation perspective.

Our conceptual framework (Fig. 1) draws upon Stremersch et al.'s universalist and social constructivist perspectives and integrates it with the notion of citability, a latent construct we develop using the literature on relationship marketing, brand equity, and research on the Matthew effect. As Fig. 1 indicates, we suggest that the impact of an article is driven by an unobserved time-varying construct "article citability" which determines the latent citation score of the article. When the citation score is positive, the article will receive a positive number of citations; otherwise, it will obtain zero citations. Our framework is consistent with the literature in that citability is informed by (1) journal effects, (2) author effects, and (3) article effects, which are all moderated by the latent article citability.

Article citability is a dynamic construct that captures the changing relationship of an article to a field. As Baldi (1998) notes, academic knowledge is cumulative and the references in a journal article highlight the relationship of an article to previous studies (Sivadas and Johnson 2005). Samiee and Chabowski (2012, p. 368) point out that "articles influence a field only if they are heavily cited by others." Our conceptualization of citability is inspired by the extensive work in the area of relationship marketing (cf. Dwyer et al. 1987; Morgan and Hunt 1994). Berry (1983, p. 25) noted that relationship marketing is all about a process of "attracting, maintaining, and enhancing customer relationship." In a similar vein, we suggest that an article's citability is conditioned on its ability to draw and maintain attention amongst the community of scholars and embed itself in a relational network. As Morgan and Hunt (1994, p. 22) note, relationship marketing refers to "activities directed toward establishing, developing, and maintaining successful relational exchanges." When a paper is cited, the citer is acknowledging a debt of intellectual influence, whilst those cited get recognition in exchange (Garfield 1979). This subsumes notions of trust and interdependence, two concepts central to conceptualization of relationships. When one cites another, one indicates a certain level of trust in their ideas and the quality of their reasoning or conclusions.

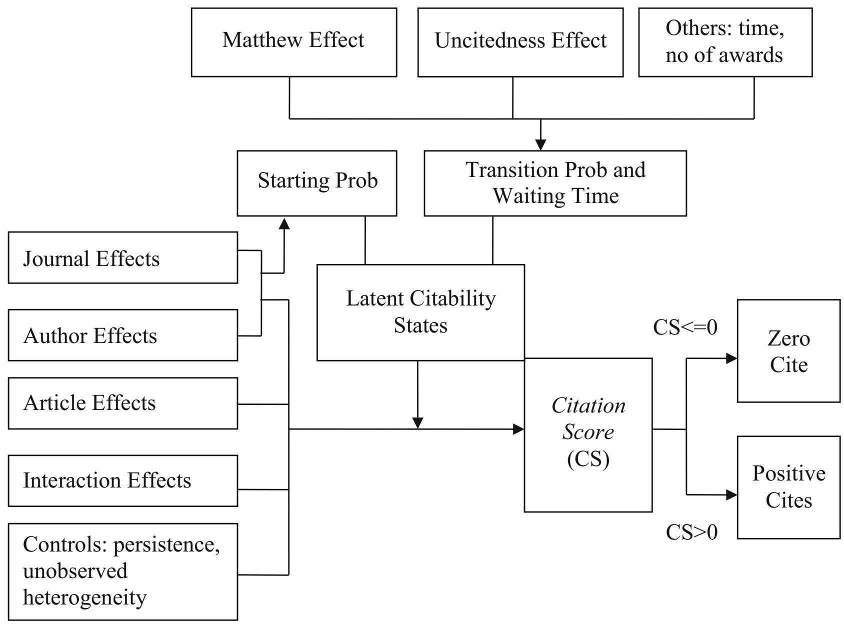
Also, as Dwyer et al. (1987) and others have noted, relationships lie on a transactional/relational continuum with a transactional or a relational orientation. Similarly, citation practices or the relationship of individual scholars with an individual paper can thus lie on such a continuum, but the collective citations reflect the relationship of the article to the discipline. While each citation event may be construed as a separate "transaction," a series of transactions (citations) can create a relationship between the paper and the field at large (Netzer et al. 2008). Researchers have recognized that relationships go through different states and such transitions can be "triggered" by a series of discrete encounters (Li et al. 2011; Netzer et al. 2008). We propose that article citability reflects a "finite set" of relational states of a paper with a field. That relationships go through different phases or states is well documented in the literature (Luo and Kumar 2013). The hidden Markov process we utilize can capture transitions between relational states and the inherently dynamic nature of relationships. Some papers never resonate with a discipline and are never able to develop a relationship; they start out and remain obscure. Others may start out slow but eventually embed themselves in a relational network, whilst others may start out strong and their relationship with a discipline may strengthen or weaken over time. Fundamentally, relationships are ever-changing and dynamic (Dwyer et al. 1987; Luo and Kumar 2013).

The latent citability of an article is also in part the aggregate-level average assessment of quality by the scholars in a given field which is subsumed in the relational state of a paper with the field. Initial quality judgments are made by journal reviewers and editors during the review process. Then, once published, quality assessments may initially typically be made based on the perceived quality of the journal the article is published in or author reputation.

Our goal is to capture the influence dynamics and build a model that can better explain the influence of an article. When a scholar reads an article, he/she may form his/her perceptions about quality of the work, which will determine whether he/she cites the article in his/her own work since citations are a form of acknowledging intellectual inspiration (Cronin 1981; Garfield 1979). Sometimes a scholar may not read the article before citing it and may simply base the citation decision on the quality perception he/she obtains indirectly from other studies that cite the original article. Thus the relational embededness of an article assumes greater importance. As time goes by, more new information about the published article may emerge. For instance, the article may win some prestigious awards, become famous over time, or become obsolete over time. Therefore, a discipline's relationship with an article may move from state to state and dynamically change by going up or down depending on the new information received. Our focus is not on individual scholars' citation decisions. We suggest that the main effects such as article, author, or journal effects are affected by the field's aggregate article citability state. Additionally, a field's relationship with an article can change dynamically over time from high to low or low to high given the changes in the status of the paper over time.

We measure citability through a variety of known (citations, author standing, journal reputation) and latent variables. We define article citability as the relational state of a paper with the field. Such a relational state in part captures article "brand equity" and will be driven by the visibility, quality, and ability to elicit the correct response amongst readers (Keller 2012). We suggest that by understanding article citability we can explain the extent to which papers become influential and capture the ever-evolving levels of influence. We explicitly account for the dynamics and the latent citability states of articles using a hidden Markov model (HMM) and allow the journal, author, and article effects to dynamically change over time according to their citability states. We also allow the starting probability of the HMM states to be affected by the author effects and journal effects in the publication year, and the transition probability and waiting time of the HMM to be driven by the article-level Matthew effect, uncitedness effect, and other time-varying variables. We next discuss several variables in detail.

Fig. 1 Conceptual framework



Journal effects

Much of the research on scientometrics has focused on journal-level impact and journal rankings. As Baumgartner and Pieters (2003) have noted, the perceived quality and status of the journal significantly influences an article's impact. Scholars have to make judgment calls about what to read in order to use time efficiently, and consequently work published in higher status journals is likely to draw greater attention. More prestigious journals are also more accessible and widely available (Sivadas and Johnson 2005). Furthermore, as article citability captures a paper's current and changing relationship to a field and reflects global assessment and brand equity of the paper, the impact of the journal status may depend on the unobserved article citability state such that high citability state of the article enhances the journal effects as opposed to if the article were to be in a low citability state (Netzer et al. 2008; Li et al. 2011). To put it differently, the citation score of a paper is driven by both its citability and its journal effects. Enhanced citability of an article also magnifies the journal effects. While an article in a low citability state also benefits from stronger journal effects, the effects are muted because of the inherently low citability state of the article. When that same article moves into a high citability state, the journal effects also give it a greater boost.

Author effects

Our author effects dimension draws upon Stremersch et al.'s social constructivist perspective. This essentially taps into the halo the work of famous scholars can have, i.e., the fact that their work is more likely to be noticed and cited (Merton 1968; Cole and Cole 1973; Medoff 2006). This greater than proportional (proportional to article quality or contribution) citation of the work of well-known scholars is commonly called the (author-level) Matthew effect (Merton 1968). The phrase "Matthew effect" is derived from the Gospel of Matthew (25:29), "for unto every one that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath." The Stremersch et al. conception of the Matthew effect was limited to famous authors. Merton (1968) suggested that better regarded scholars get greater recognition for equivalent quality work as compared to less known or less established scholars. As Tol (2009, p. 420) puts it, "famous works are more easily noted, and authority lends weight to an argument." In addition, the impact of the author reputation may also depend on the unobserved article citability state such that high citability of the article may strengthen the author effects. We argue that a paper needs to be in a high citability state for it to have greater traction, and articles that have such traction benefit more from author reputation. The effect of author reputation is muted or considerably diluted when the article is in a low citability state.

Article effects

Most research on scientometrics has focused on journal-level effects rather than individual article-level effects. As Chow et al. (2007) suggest, articles should be evaluated on their own merit and publication in a top journal is not necessarily a good proxy for the quality of the article. Similarly the subject area of the article has been shown to influence impact, with certain subject areas such as relationship marketing and services marketing tending to be more cited than other subject areas such as sales (Stremersch et al. 2007; Bettencourt and Houston 2001). Subject area and article quality fall within what Stremersch et al. identified as the universalist perspective on citations, and contribute to overall article citability. Thus, similar to the journal and author effects, the extent of article effects may also depend on the unobserved article citability state such that high citability of the article may strengthen the article effects. We measure article effect through article awards and subject area. We suggest that articles that are in the higher citability states benefit more from awards and from relating to "hotter" subject areas than articles that are in low citability states. Thus, articles in higher citability states benefit more from article effects than do articles in lower citability states.

As shown in Fig. 1, we also incorporate the interaction between the journal effects and author effects to see if author fame is less important when a paper is published in an A-level marketing journal and how this interaction effect is moderated by article citability. In addition, we control for the persistence effect of previous number of citations on current citations as well as unobserved article heterogeneity. It is important to note that although journal effects, article effects, and author effects on citation have been studied in prior research (Stremersch et al. 2007), it has all been in cross sectional studies (i.e., across articles) in a static setting. This is different from our examination of these effects on the dynamics of article citations (i.e., both within and across articles) that allows these effects to dynamically change over time according to the latent citability state in which the articles are. Given the latent nature of article citability, a discrete-state HMM is appropriate to capture the citability states since it has been employed to successfully model competitive promotions, customers' unobserved life stages, and relationship states in the marketing literature (Du and Kamakura 2006; Moon et al. 2007; Netzer et al. 2008; Li et al. 2011).

Starting probability of the HMM

When an article is in its first publication year, the scholars' initial citability state is likely to be driven by the author and journal effects, which may serve as quality signals (Spence 1973). However, the subject area of the article is unlikely to serve as a signal since anyone can choose a topic to work on. Therefore, we allow the author and journal effects but not article effects to affect the starting probabilities of the HMM as shown in Fig. 1.

Transition probability and waiting time of the HMM

After the publication year of the article, as time goes by, the article may win some prestigious awards, more scholars have access to it, it could accumulate some fame, or it may go uncited for a long time. Therefore, we allow these time-varying variables to affect the transition probabilities and waiting time of the article citability states in the HMM as shown in Fig. 1. The transition probabilities will govern which HMM state the article will jump into from its current state if a jump occurs, and the waiting time of the latent citability states determines how long the article is going to stay in the state it is in.

Article-level Matthew effect on the transition probability and waiting time of the HMM

The Matthew effect, while mostly conceptualized as an author effect (Stremersch et al. 2007), also exists for institutions (Medoff 2006), for countries (Bonitz et al. 1997), and for famous papers as well (Tol 2009). Some papers authored by less known scholars may initially draw attention for their content or for their novelty (the universalist perspective on citations) and over time become famous papers. and as Tol (2009, p. 423) points out, "often-cited papers are cited more often." So, we suggest that the Matthew effect is not simply a famous author effect, but it holds true for famous papers as well, i.e., citations beget more citations. The Matthew effect can come into effect for famous papers (written by not-so-prominent authors) or for papers published by less known authors with prestigious institutional affiliation (Medoff 2006). Thus, in our framework we separate the article-level Matthew effect from mere author prominence and incorporate the former into the HMM process. For the article-level Matthew effect, we propose a new behavioral rationale on how it works. Specifically, we posit that the Matthew effect strengthens the relationship of an article with a discipline by increasing the transition probability of the article jumping from a lower citability state to a higher state, and/or decreasing the article's time staying in the lower state and increasing its time in the higher state. These in turn result in more citations.

Uncitedness effect on the transition probability and waiting time of the HMM

Weale et al. (2004) suggest that the rate of non-citations to articles published in journals can also be used as a measure of (lack of) journal quality. Hendrix (2008) views the percentage of non-cited articles as one metric for assessing research quality. Garfield (2005) estimates that nearly half the articles indexed in ISI's database have zero citations. As Stern (1990, p. 193) put it, "in the absence of citations there is no firm and readily available evidence that a publication has contributed to the advancement of scholarship." Bibliometric scholars have examined various factors that may contribute to uncitedness such as the influence of the disciplinary citation practices of number of references per article or the journal in which an article was published (Seglen 1992; Stern 1990). Seglen (1992) and others have noticed the practice of concentration of citations with a small fraction of articles accounting for nearly half of the citations accrued by a journal.

Different from these studies, our focus is on the issue of what happens to a paper's long-term impact if it does not get cited early on. We focus on short-term uncitedness here. This is to be distinguished from long-term uncitedness, which is defined by

Stern (1990) as papers that go uncited eight years after publication. A paper that may not draw citations early on (within the first two years) may get some citations over the first 8 years; however, we focus on the consequences of not generating citations early on and suggest that not generating citations early on has a detrimental effect on the long-term impact of an article. Consequently, zero-cited articles become less and less likely to get cited over time. We term this the *uncitedness effect.* As Tol (2009) points out, citations beget more citations. When an author cites a scholar, they draw attention to that scholar's work. Thus a citation can serve as an advertisement or a signal of quality (Spence 1973). It may also serve the function of slowing down the obsolescence of published work by keeping it fresh in the mind of scholars working in the area, as when they read more recent articles in the area they are referred to older articles that are cited therein. Our model (unlike prior models) can explicitly account for zero citations and estimate the effect of early uncitedness on the rate of later citations. Similar to the article-level Matthew effect, we expect that early uncitedness decreases citability of the article by increasing the transition probability of the article jumping from higher citability state to lower state, and/or by increasing the article's time staying in the lower state and decreasing its time in the higher state. These in turn result in lower citability and less citations over time.

Data

Traditionally, the *Journal of Marketing*, *Journal of Marketing Research*, and *Journal of Consumer Research* are ranked as the major journals in the field, with *Marketing Science* as a more recent addition to this group (Lehmann 2005; Seggie and Griffith 2009). This study also includes other recognized journals with a broad rather than specialized coverage of marketing topics namely, the *Journal of the Academy of Marketing Science*, *International Journal of Research in Marketing*, and *Journal of Retailing.* Various studies suggest these seven journals are ranked in the top ten in an international context (Mort et al. 2004; Guidry et al. 2004). In the business disciplines, journals with a more general focus tend to have higher citations than do specialized journals (Zinkhan and Leigh 1999).

These journals were also selected for our study because of their breadth, rather than specialization, in marketing topics. Davis (1998) describes a number of difficulties entailed in making "apples to oranges" comparisons between the general interest and the specialized journals with respect to journal citation impact. While it may not be universally viewed as a general marketing journal, the *Journal of Retailing* is included in this group of general marketing journals rather than as a specialized journal because it has a broad substantive and methodological focus that includes services marketing, econometric models, advertising, sales promotions, supply chain management, and consumer behavior. Additionally, in terms of structural influence of journal intercitations, Baumgartner and Pieters (2003) rate *JR* as a core marketing journal rather than a marketing application journal. The other six journals, with the exception of *JCR,* also fall into the sub-area of core marketing journals as opposed to managerial, application, and education journals in Baumgartner and Pieters' study.

This study analyzes a total of 1,591 articles consisting of all articles published during the calendar years 1996–2003 for the seven selected journals. For this study, article citation data were downloaded from Scopus, a citation tracking service developed by Elsevier as a competitor to Thomson ISI (i.e., data source used in previous studies; see Stremersch et al. 2007). Elsevier began the development of Scopus in 2002 and launched the product in 2004 after extensive testing. Scopus has worked backwards to create coverage of social science journals to 1996, which is the starting point of this study.

Scopus does not include journals from the arts and humanities (journals that seldom cite marketing articles), but the coverage of the sciences, social sciences, and business is greater than ISI. Scopus has a current database of over 21,000 titles from 5,000 publishers including 20,000 peer-reviewed journals, and conference proceedings, trade publications, and book series. Scopus is also regarded as having greater international coverage than Thomson ISI. Consequently, Scopus may record more citations than ISI for these marketing articles. For example, Zinkhan (2005) lists the 15 most cited *JAMS* articles published between 1998 and 2004. Zinkhan reports these articles have a total of 416 citations in the ISI database as of December 2004; in the Scopus database these same articles have a total of 511 citations as of 2004, representing an increase of 22.8%. The higher level of Scopus citations is consistent across all of the articles with one exception. The reason is simple. ISI is based on the premise of "scientometry," that useful knowledge in any given discipline is only contained in a small subset of journals. Thus ISI tracks *only* a smaller subset of journals in each discipline. ISI tracked citations to only 20 marketing journals in 2004, and in 2012 it tracked 30 marketing journals. Thus, ISI does not record citations (i.e., examine and record references) to articles published in other marketing journals besides the small number of journals they track. Papers cited in articles published in marketing journals such as *Journal of Personal Selling & Sales Management*, *Journal of Product & Brand Management*, and *Journal of Consumer Marketing* (to name a few) are not recorded. Thus, ISI undercounts the number of citations that accrue to specific articles in comparison to Scopus.

A second problem that led us to use the more comprehensive Scopus database is that ISI has only recently started tracking citations for some of the respected marketing journals in our list such as *Journal of the Academy of Marketing Science* and *International Journal of Research in Marketing*, and data do not go back to the mid-1990s for these journals. Sivadas and

Johnson (2005) suggest that it takes up to 6 years post-publication for citations to a specific article to peak, and we wanted to allow a sufficient timeframe from publication to data downloading. Thus, for these two reasons we choose the more comprehensive Scopus. We would like to emphasize that Scopus tracks all the marketing journals tracked by ISI plus several more and is thus more comprehensive. Furthermore, Archambault et al. (2009) find there is a very high degree of consistency between Scopus and ISI's SSCI databases.

Scopus and ISI also face competition from the citation tracking capability of Google Scholar, which is a freely available internet product. A large number of academic publishers provide Google Scholar with article abstracts and reference lists (but not full-text). Consequently Google Scholar may record even more citations for these marketing articles. However, Google scholar may also record citations to unpublished articles as well that are posted on individuals' websites. Hence, we chose to go with Scopus.

We downloaded citations at the end of 2006. Thus we allowed for a minimum of three years post-publication and a maximum of 10 years from date of publication to date of citation count download. As Katerratanakul et al. (2003) point out, it typically takes about 2 years from date of publication for citations to start accruing. For this study, year-by-year citation data were downloaded for each of the 1,591 articles published by the seven journals during the period of 1996–2003. Annual citations for all articles were downloaded through the end of 2006.

We cross-checked the Scopus database against the table of contents for the seven journals and found only one substantial misclassification—the entire March 2003 issue of *JCR* was mistakenly coded as 2002. This was corrected. We also did not include five erratum pages from *MKS* in our analysis.

We found that there exist big citation differences across the seven journals. *JM*, *JMR*, and *JCR* have the highest mean or median number of citations, highest percentage of articles with 20 or more cites, and lowest percentage of zero-cited articles. Further, different articles seem to have different citation dynamics. The articles with high total citations at the end of the data period tended to grow much faster over time than those with low total citations (see Appendix A).

Model setup

To capture article-level citability dynamics, we adopt a modified version of the count model—a discretized Type I Tobit model with a hidden Markov process proposed by Li et al. (2005). This model has several advantages over the commonly used NBD model for counting data in our context (Burrell 2003; Mingers and Burrell 2006; Stremersch et al. 2007). First, the latent construct in the Tobit model is suitable for capturing the unobserved dynamic citation score and can easily accommodate the various effects on citation score in our conceptual framework in Fig. 1. Second, the Tobit model can explicitly account for the dynamics of zero citation and positive citation over time. Third, the hidden Markov process can capture the latent citability states among scholars and their dynamic changes over time. Although past studies (Stremersch et al. 2007) have included time as a covariate in the NBD model to capture the time impact, the parameters (i.e., the author, journal, and article effects) in their models are all static and do not change over time as we allow in the HMM process. Given the count nature of article-level citations, we discretized a Type I Tobit model and take log transformation of the number of new citations at time t to capture the long tail. We also compare the proposed model to the NBD model in the empirical analysis section and show that the former outperforms the latter. The modification from Li et al. (2005) is that while the HMM is homogenous in their study, we allow the HMM to be heterogeneous and be affected by the Matthew effect, uncitedness effect and other time-varying factors, which will shed light on the dynamics of article citations over time. We model the number of new citations (Z_{it}) for article i at time t as follows

$$Z_{it} = \begin{cases} k = Floor\left(\exp\left(Z^*_{it}\right)\right) & \text{if } Z^*_{it} > 0 \text{ and } \ln k \le Z^*_{it} < \ln(k+1) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

$$Z^*_{it} = \gamma'_{is} X_{it} + \varepsilon_{its}, \varepsilon_{its} \sim N\left(0,\ \sigma^2_s\right), \text{with probability } p_s \text{ such that} \sum_{s=1}^{S} p_s = 1 \tag{2}$$

where the article's latent citability state is indexed by s, which we explain in the next sub-section. Z^*_{it} is a latent dynamic variable that captures the article's citation score, k is a positive integer, and *Floor*(Y) is the integer component of Y where the discretization occurs. As depicted in the conceptual framework, there is a threshold of zero for an article to ever get a positive cite. That is, when its citation count is greater than zero (Z^*_{it}>0), the article will obtain positive cites (Z_{it}) and zero cite otherwise. X_{it} includes the article, author, journal effects, and the interaction effects shown in the conceptual framework. We also take log transformation for the dependent variable Z_{it} (i.e., the exp() operator in Eq. 1). γ_{is} (including an intercept) is a vector of article- and citability-state-specific parameters to be estimated and captures the impact of the various effects. σ^2_s is the latent citability state specific variance of unobserved disturbance in citation score to be estimated as well.

An article's latent citability state

The parameters of our discretized Tobit model (Eq. 2) are indexed by state s at each time period. This state is meant to capture an article's latent citability which governs the dynamic evolvement of its citations over time. We assume an article can be allocated to one of S latent states at each time period and the total number of states S will be determined empirically. The transition among these states is governed by a first-order continuous-time discrete-state hidden Markov model (HMM) (Montgomery et al. 2004; Li et al. 2011). For brevity we interpret our latent states as an indicator of the article's citability state among scholars. Our interpretation of states is based upon a comparison of the estimated coefficients different across states and summary statistics. However, our interpretation and labeling of citability states are not unique, just as a label for a segment in cluster analysis or factor in factor analysis is not unique (Li et al. 2011).

A hidden Markov model of latent citability states

We use an S x S matrix M_{it} to denote the probabilities for article i to transition to another state at time t:

$$M_{it} = \begin{bmatrix} 0 & P_{it12} & \cdots & P_{it1S} \\ P_{it21} & 0 & \cdots & P_{it2S} \\ \vdots & \vdots & \ddots & \vdots \\ P_{itS1} & P_{itS2} & \cdots & 0 \end{bmatrix}. \tag{3}$$

Each element in the transition matrix P_{itmn} represents article i's probability of transiting from state m at t-1 to state n at time t. Hence, $0 \leq P_{itmn} \leq 1$, and the row sum is one.

The diagonal elements of M_{it} are zeroes since we do not allow same-state transitions. Instead, we capture persistence within a state as a waiting time for the state, which is the duration an article stays in one particular state. We define $W_{it}(s)$ as the waiting time in state s and assume it follows a gamma distribution in a continuous time domain (Li et al. 2011):

$$\Pr(W_{it}(s)|\lambda_{it}(s), k_i(s)) = \frac{k_i(s)^{\lambda_{it}(s)}}{\Gamma(\lambda_{it}(s))} W_{it}(s)^{\lambda_{it}(s)-1} e^{-W_{it}(s)k_i(s)}. \tag{4}$$

$\lambda_{it}(s)$ is the shape parameter and $k_i(s)$ is the inverse scale parameter for state s. Notice if $\lambda_{it}(s)=1$ we have an exponential distribution. Being article specific, $\lambda_{it}(s)$ and $k_i(s)$ determine how long article i stays in state s. More specifically, the expected waiting time until the next state equals the ratio of the shape parameter to the inverse scale parameter:

$$E[W_{it}(s)] = \frac{\lambda_{it}(s)}{k_i(s)}. \tag{5}$$

Unlike the homogeneous HMM Du and Kamakura (2006), Li et al. (2005) and Moon et al. (2007) use, we adopt a heterogeneous HMM and allow the article's waiting time (e.g., the shape parameter $\lambda_{it}(s)$ in Eq. 4) to be affected by the article's award status, time impact, the Matthew effect, and the uncitedness effect. Specifically, we assume $\lambda_{it}(s)$ follows a log-normal distribution:

$$\log(\lambda_{it}(s)) \sim N\left(\overline{\lambda}_{it}(s), \sigma^2_\lambda\right), \tag{6}$$

where its mean $\overline{\lambda}_{it}(s)$ is a function of the article's award status, time impact, the Matthew effect, and the uncitedness effect:

$$\begin{aligned} \overline{\lambda}_{it}(s) = \alpha_{0is} &+ \alpha_{1is} Time_{it} + \alpha_{2is} Time^2_{it} + \alpha_{3is} Award_{it} \\ &+ \alpha_{4is} \log(rank)_{it-1} + \alpha_{5is} \log^2(rank)_{it-1} \\ &+ \alpha_{6is} \log(years\_nocite)_{it-1} \\ &+ \alpha_{7is} \log^2(years\_nocite)_{it-1} \end{aligned} \tag{7}$$

The coefficient α_{0is} captures an article's intrinsic tendency to stay in citability state s. *Time* denotes the number of years passed. α_{1is} and α_{2is} capture the time impact and its potential non-linear effect on waiting time, respectively. *Award* refers to the number of major journal awards won by the article and α_{3is} capture the award effect. log(*rank*) and log(*years_nocite*) refer to the logarithm of the citation ranking of the article and the cumulative number of years with zero cites up to the end of previous year. α_{4is} and α_{6is} capture the main impact of these two variables on waiting time, while α_{5is} and α_{7is} denote the Matthew effect and uncitedness effect, respectively.

Initial citability state probabilities of the hidden Markov model

We define the initial state probabilities of article i residing in state s for $s = 1,\ldots, S$ at time 0 as a vector $\prod_i = (\pi_i(1), \ldots, \pi_i(S))'$. The row vectors of the transition matrix and the vector of initial starting probabilities are assumed to follow a Dirichlet distribution:

$$\mathbf{P}_{itj} \sim D(\boldsymbol{\tau}_{itj}), \quad \Pi_i \sim D(\eta_{is}), \tag{8}$$

Where $\mathbf{P}_{itj}$ denotes the j^{th} row of the transition matrix $\mathbf{P}_{it}$, and τ_{itj} and η_{is} refer to the hyper-parameters for the transition and starting probabilities, respectively. Similar to the specification of the waiting time intensity, we assume τ_{itj} and η_{is} follow a log-normal distribution:

$$\log(\tau_{itj}) \sim N(\overline{\tau}_{itj}, \sigma^2_\tau), \quad \log(\eta_{is}) \sim N\left(\overline{\eta}_{is}, \sigma^2_\eta\right). \tag{9}$$

In order to take into account the impact of the author and journal effects on an article's starting probabilities in state s, we define $\overline{\eta}$ as a function of an article's author effect, journal effect, and article effect at time 0 (e.g., year of article publication). That is,

$$\overline{\eta}_{is} = \omega_{0is} + \omega_{1is} \cdot V_{i0}, \tag{10}$$

where V_{i0} consists of the author and journal effects as in Eq. 2 at time 0. Coefficients ω_{1is} measure how these variables at time 0 affect the probability that an article starts in state s.

Article heterogeneity and estimation

We model the article-level unobserved heterogeneity following a random-coefficient approach in a hierarchical Bayes framework (Rossi et al. 1996). That is,

$$\gamma_{is} \sim MVN\left(\overline{\gamma}_s,\ \Lambda_s\right) \tag{11}$$

Given the model specification above, we have the following conditional likelihood function for article i at time t:

$$\Pr(Z_{it}|s) = \int_{\gamma_{is}} (\Pr(Z_{it} = k|s))^{I(Z_{it}>0)} (\Pr(Z_{it} = 0|s))^{I(Z_{it}=0)} f(\gamma_{is}) d\gamma_{is} \tag{12}$$

where $\Pr(Z_{it}=k|s)=TN_{[\ln k, \ln(k+1))}[\gamma_{is}^{'}X_{it}, \sigma_s^2]$ and $\Pr(Z_{it}=0|s)=TN_{(-\infty,0]}[\gamma_{is}^{'}X_{it}, \sigma_s^2]$. $I()$ is an indicator function, and $TN_{(a,\ b)}$ denotes the truncated normal distribution between values a and b. $f(\gamma_{is})$ is the heterogeneity distribution specified in Eq. 11. The unconditional likelihood and identification of the hidden states are presented in Appendix B.

Given the high-dimensional integrals in the likelihood function, a Hierarchical Bayes approach is demonstrated to be a good choice for estimation (Rossi et al. 1996). We use the Gibbs Sampler and the Slice Sampler (for truncated Normal distributions; see Damien et al. 1999) to obtain draws from the full conditional distributions of the parameters (Chib and Greenberg 1995). Additionally, using the Data Augmentation approach (Tanner and Wong 1987), we treat the unknown utilities Z_{it}^* as parameters and make draws for them from their own full conditional distributions. We estimate the empirical model using a program coded in “C++”. The chain for the Gibbs sampler was run for a total of 50,000 iterations. The first 40,000 iterations were discarded as “burn-in” before convergence was attained (Gelfand and Smith 1990). The remaining draws were used for inference.

Empirical analysis and results

Measures

We randomly divide the total sample into two parts: Part I with three quarters of the data (1,193 articles and 8,851 observations) constitutes the estimation sample, while the remaining quarter forms the holdout sample (398 articles and 2,961 observations) for model comparison purpose. The simple statistics are presented in Table 1.

Dependent variable Since we are interested in article citability and the dynamics of the citation, we operationalize the dependent variable as the number of new citations that the article receives in year t. Citations are shown to be an objective measure of influence, impact, or attention (Pieters and Baumgartner 2002). As seen in Table 1, the annual number of citations per article is 3.27 with median 2 and a standard deviation of 0.05.

Independent variable: article-level Matthew effect Merton (1968) defines the Matthew effect as the phenomenon that fame breeds fame in the form of citations. To capture the potential article-level Matthew effect, we adopt the approach proposed by Tol (2009) on the basis of the theory of growth of firms such that the log of the firm size is proportional to the log of the rank of the firm (Gibrat’s law) (Ijiri and Simon 1974; Simon 1955). Tol (2009) uses the following empirical test: ln(citations) = α + $\beta\times$ln(rank) + $\gamma\times\ln^2$(rank), where “citations” is the total number of citations an article receives at a particular time and “rank” is the ranking of an article (i.e., article with most previous citations = rank 1) based on the number of the previous citations to that article. In this equation, γ is used to measure the article-level Matthew effect (second-order effect) such that if γ is significantly less than zero, there are increasing returns to scale and hence a Matthew effect (i.e., it is due to the fact that the ranking is reverse-scored such that the article with most previous citations = rank 1). Otherwise, there is no Matthew effect.

In our study, we also capture the potential article-level Matthew effect through the coefficient of the square of log of article rank based on the cumulative citations of all papers in the sample in previous year.[1] To capture the first-order effect of article rank, we also incorporate the variable ln(rank) as an independent variable. When we rank the articles in the sample, we exclude those with no citations in the previous year. In other words, the value of this variable will be zero for those articles without any citations. There are two reasons why we do that. First, it is consistent with the definition of the Matthew effect such that often-cited papers get cited more often. If an article has not been cited at all and hence has no recognition, the Matthew effect does not apply. Second, we are also interested in the potential existence of uncitedness effect such that zero-cited papers become less and less likely to be ever cited over time. Therefore, we separate articles with positive citations from those without citations to identify the Matthew effect and uncitedness effect respectively. From Table 1, we can see that the average log rank of an article with positive cites is 3.64 with standard deviation 0.78.

Stremersch et al. (2007) measured the author-level Matthew Effect by looking at authors’ editorial board membership, number of publications in five journals, and ranking of the business school they were affiliated with. We capture editorial board membership in our author effect variable. One strength of our model is that our ability to capture unobserved variables (see below) makes it more easily implementable.

[1] We also tried the square of article rank without log transformation. The results remain the same but with less model fit.

Table 1 Simple statistics

Effects	Variables	Mean	Std. Dev.	Min	Median	Max
Dependent variable	No of new cites	3.27	0.05	0	2	109
Author effects	No of authors	2.30	0.93	1	2	8
	Editbrdjm	0.35	0.58	0	0	3
	Editbrdjmr	0.30	0.53	0	0	3
	Editbrdjcr	0.28	0.55	0	0	4
	Editbrdmks	0.24	0.53	0	0	3
Journal effects	JM	0.16	0.36	0	0	1
	IJRM	0.12	0.32	0	0	1
	MKS	0.12	0.32	0	0	1
	JMR	0.19	0.39	0	0	1
	JAMS	0.14	0.35	0	0	1
	JCR	0.16	0.37	0	0	1
Article effects	No of awards	0.06	0.26	0	0	2
	Sub_newprod	0.06	0.25	0	0	1
	Sub_b2b	0.06	0.24	0	0	1
	Sub_relationship	0.07	0.26	0	0	1
	Sub_prodbrand	0.12	0.32	0	0	1
	Sub_ad	0.06	0.25	0	0	1
	Sub_pricing	0.08	0.27	0	0	1
	Sub_promotion	0.06	0.23	0	0	1
	Sub_retailing	0.10	0.30	0	0	1
	Sub_strategy	0.09	0.29	0	0	1
	Sub_sales	0.04	0.20	0	0	1
	Sub_method	0.16	0.37	0	0	1
	Sub_services	0.08	0.28	0	0	1
	Sub_cb	0.17	0.37	0	0	1
	Sub_international	0.05	0.23	0	0	1
	Sub_ecom	0.03	0.17	0	0	1
Matthew effect	Log(Rank last year)	3.64	0.78	0	3.81	4.58
Uncitedness effect	Log(no of years with zero cites)	0.64	0.49	0	0.69	2.30

Independent variable: uncitedness effect As discussed in the conceptual framework section, we want to empirically test for the existence of uncitedness effect. Similar to the operationalization of the article-level Matthew effect, we construct an independent variable of log of the number of years with zero citations for those articles that never got cited (first-order effect) as well as its square term (second-order effect). Note that the value of these variables will be zero for those articles with positive citations. Thus, the coefficient of the square term can be used to test the uncitedness effect. In the sample, the log of number of years with zero citations is 0.64 on average with standard deviation 0.49.

Independent variables: article effects author effects, and journal effects We capture the article effects on paper citation through two measures: number of journal awards received and subject area of the article. First, the awards we consider include the following best article awards at various journals: Best Article Award (*IJRM*), Best Article Award (*JCR*), Harold H. Maynard Award (*JM*), MSI/H. Paul Root Award (*JM*), Paul E. Green Award (*JMR*), William F. O'Dell Award (*JMR*), and John D.C. Little Award (*MKS*), Sheth Foundation best paper award (*JAMS*), best article award (*JR*). Since these awards are chosen by editorial boards of the corresponding journal and may be considered the choice of the highest-quality article by leading scholars in marketing, the number of journal awards received should be a good indicator of article quality. The average number of awards received in our sample is 0.06 with standard deviation 0.26 and the maximum of two awards. Second, we adopt the categorization scheme of subject areas used by Stremersch et al. (2007) and come up with 16 subject areas (therefore 15 dummy variables as the last one "other" as the default case) as shown in Table 1. The subject area is the subject on which an article focuses, and an article may belong to multiple subject areas. Stremersch et al. (2007) have shown that the categorization is reliable and the subject areas have important impact on article citations. The subject areas of the articles in our sample are fairly broad with slightly more articles on consumer behavior, methodology, product and brand management, retailing, and strategy.

Two measures are used to capture the author effects of an article on its citations. The first one is number of authors for a particular article. We can see that in the sample the average number of authors per article is 2.3 with minimum 1 and maximum 8. The second measure is the number of editorial board membership among the authors for a particular article for the four top marketing journals: *JM*, *JMR*, *JCR*, and *MKS*. Table 1 shows that the mean of the number of editorial board membership per article for the four journals is 0.35, 0.30, 0.28, and 0.24, respectively.

We incorporate the journal effects using six journal dummy variables with *JR* as the default journal. In the sample, the frequency of articles across these journals is roughly the same with slightly higher frequency for *JMR*, *JM*, and *JCR*.

Independent variables: interaction effects We are interested in the interaction between author effects and journal effects to see if author fame is less important when a paper is published in an A-level marketing journal. To simplify the analysis, we use the following variables to capture the author effects or journal effects: the total number of editorial board memberships across the four top marketing journals—*JM*, *JMR*, *JCR*, and *MKS* (author effect)—and whether it is an A-level journal—*JM*, *JMR*, *JCR*, and *MKS* (journal effect).

Independent variables: controls To control for the time impact on article citations, we include a lag log of cumulative citations as a covariate. This variable may capture the persistence of the article citations (Seetharaman 2004). To control for any other missing article-specific variables, we incorporate it through the article-specific intercept. Unobserved article heterogeneity is accounted for using the random-coefficient approach in the heterogeneity equation (Eq. 11). We also check the correlations among the covariates and find that multicollinearity is not an issue in this study.

Model comparison

To validate the proposed modeling framework, we compare it to several benchmark models. This first benchmark is the aggregate NBD model with the same covariates of the proposed model but without accounting for article heterogeneity. The second benchmark model is the heterogeneous NBD model, which adds article heterogeneity to the first model, both of which are non-nested within the proposed model. Both NBD model setups are standard and available upon request from the authors. The other models are nested benchmark models of the proposed model. We estimate the aggregate version of the proposed model without article heterogeneity, and the proposed model with HMM of one state, two states, three states, and four states, respectively, and report the results in Table 2. To test if allowing the citability states switch up only in the HMM fits the data better, we also estimated the proposed model with HMM switch up only of one state, two states, three states and four states, respectively. Due to space constraint, we report only the best model out of the four versions—the two-state discretized model with switch up only in HMM. The log marginal density (Chib 1995), mean absolute error (MAE), and square root of mean square error (RMSE) for each of the seven models for both the estimation sample and holdout sample are presented in Table 2.

The results show that the heterogeneous NBD model outperforms the aggregate NBD model with higher log marginal density and lower MAE and RMSE for both the estimation sample and holdout sample, demonstrating the importance of accounting for article heterogeneity. The aggregate discretized Tobit model outperforms the aggregate NBD model, and the one-state discretized Tobit model (i.e., the heterogeneous discretized Tobit model) outperforms the heterogeneous NBD model, indicating the appropriateness of the proposed discretized Tobit framework. Out of the four versions (one state, two state, three state, and four state) of the proposed model with HMM, the two-state discretized Tobit is the best model in terms of higher log marginal density, and lower MAE and RMSE for both the estimation sample and holdout sample. This model also outperforms the best model among the four versions of the proposed model with HMM with switch up only—the two-state discretized model with switch up only in HMM, demonstrating the possibility of HMM states switching both up and down over time. Since the two-state proposed model is the best model overall, we will focus on the discussion of the estimation results of this model. Also, to show the potential estimation bias in the heterogeneous discretized Tobit model without HMM, we also report its estimation results in Table 3 (i.e., the results for the One-State Model in the table). The significant estimates (i.e., zero does not lie in the 95% posterior probability interval of the estimate) are highlighted in bold.

Estimation results of the two-state model

From Table 3, we can see that the intercept of State 2 is higher than that of State 1 following the identification condition of HMM, indicating articles in the second state tend to receive more new citations. Articles in State 2 also have slightly higher persistence effect than those in State 1.

Author effects, journal effects, and article effects Interestingly, we find that the author effects, journal effects, and article effects vary significantly across the HMM states. The number of authors significantly decreases the number of new citations in the second state while it has insignificant impact in State 1,

Table 2 Model comparison

Model	Estimation sample			Holdout sample	
	Log marginal density	MAE	RMSE	MAE	RMSE
Aggregate NBD model	−22812.00	11.98	31.44	16.71	32.30
Heterogeneous NBD model	−22382.30	11.38	21.77	12.26	22.88
Aggregate discretized Tobit	−22752.16	11.93	25.96	14.71	27.13
One-state discretized Tobit	−13795.06	10.94	15.34	12.12	17.86
Two-state Tobit with switch up only	−6556.74	2.72	7.25	9.88	12.91
Two-state discretized Tobit	−6614.95	2.67	6.63	9.39	11.67
Three-state discretized Tobit	−6843.09	10.61	12.53	11.98	16.82
Four-state discretized Tobit	−7162.75	11.24	15.27	12.12	21.91

possibly because it is considered as a negative signal in State 2 but not in State 1 by scholars. Articles authored by those with prestigious editorial board membership at *JM*, *JMR*, *JCR*, and *MKS* significantly attract more new citations in State 2. However, in State 1 only the editorial board membership at *JM* and *JCR* out of the four top marketing journals has significantly positive impact on the number of new citations. In the first state, the U.S.-based journals (all the journals in this study except *IJRM*) tend to attract more new citations compared to the base journal, *JR*, especially the top four marketing journals (*JM*, *JMR*, *JCR*, and *MKS*). The more internationally oriented journal *IJRM* attracts less citations compared to *JR*, in both states. However, in State 2, only *JM* and *JCR* attract more new citations compared to *JR*.

In terms of the subject areas, in the first state, we find that articles in the areas of relationship marketing, services, consumer behavior, international marketing, or ecommerce tend to be popular topics and receive more new citations, while articles on pricing, promotion, sales, or methodology attract considerably less citations. But in State 2, the story is somewhat different. In the second state, articles on new products also tend to attract more new citations, while those on business–to-business marketing, product and brand management, advertising, pricing, retailing, sales, or methodology attract considerably less citations. This may be due to the fact that the popularity of subject areas in different states is different.

Lastly, regardless of the states, surprisingly, publishing in A-level marketing journals (journal effects) tends to mitigate the positive impact of prestigious editorial board membership at *JM*, *JMR*, *JCR*, and *MKS*, with even more negative interaction effect in State 2. This indicates that author fame is less important when a paper is published in an A-level marketing journal, especially in the high citability state. The variance of new citations in State 2 seems to be larger than that in State 1, indicating more variation of citations in State 2.

Results of the HMM We find that in the year of publication, an article on average is likely to start in the first state with a much higher probability of 0.78 (0.02) compared to the probability of 0.22 (0.02) in the second state. (The numbers in parentheses are the posterior standard deviations.) The average waiting time of State 1 and 2 is 5.54 (0.06) and 5.61 (0.05) years, respectively, which indicates slightly longer duration in the higher state once in that state.

Table 4 presents the estimation results on what drives the starting probability of the HMM. To interpret the results, we need to point out that the starting probability of the HMM follows a Dirichlet distribution with its hyper-parameters as functions of covariates shown in Table 4. The impact of one covariate on the expected starting probability in one particular state will be determined by the relative ratio of the estimated coefficients for the covariate across the two states. For instance, take the intercept as an example: it has insignificant impact on the hyper-parameter in State 1 but negative impact on the hyper-parameter in State 2. This means that overall, intrinsically an article is more likely to start in the first state than in State 2. Similarly, we find that the number of authors, editorial board membership at *JMR*, or submitting to the journals *IJRM*, *MKS*, or *JMR* (compared to submission to the base journal *JR*) increases probability of the article starting in the low state—State 1. In contrast, the editorial membership at *JM* or submitting to *JCR* results in higher starting probability of State 2, possibly due to the relatively higher citability from these actions. Interestingly, we also find that editorial board membership at *JCR* or *MKS*, or submitting to the journals *JM* or *JAMS* (compared to submission to the base journal *JR*) has no significant impact on the article's starting probability in either state.

Table 5 summarizes the estimation results for the waiting time equations. Based on the estimated intercepts, we find that once in State 1, intrinsically an article is likely to stay for shorter time in the state and more likely to jump to the other state compared to in State 2, which indicates higher persistence in the high state. As time goes by, the article is less likely to stay long in both states. However, there is a U-shaped curve-linear impact of time on the waiting time in State 1 but not in State 2. This means that as time passes by and after certain point, the article in the first state tends to stay in the state longer and longer, while this trend does not exist for the

Table 3 Estimation results for the discretized Tobit Model with HMM

Effects	Variables	One-state model	Two-state model	
			State 1	State 2
	Intercept	**−0.619 (0.059)**	**−0.333 (0.155)**	−0.001 (0.181)
Persistence effect	Lag of log no of cites	**0.991 (0.001)**	**0.813 (0.001)**	**0.825 (0.001)**
Author effects	No of authors	0.002 (0.005)	−0.006 (0.004)	**−0.025 (0.005)**
	Editbrdjm	**0.104 (0.016)**	**0.082 (0.018)**	**0.096 (0.017)**
	Editbrdjmr	**0.127 (0.033)**	0.046 (0.027)	**0.129 (0.026)**
	Editbrdjcr	**0.092 (0.025)**	**0.056 (0.025)**	**0.123 (0.021)**
	Editbrdms	**0.056 (0.028)**	0.046 (0.033)	**0.063 (0.025)**
Journal effects	JM	**0.306 (0.051)**	**0.460 (0.040)**	**0.278 (0.042)**
	IJRM	**−0.257 (0.059)**	**−0.151 (0.039)**	**−0.451 (0.051)**
	MKS	**0.205 (0.059)**	**0.269 (0.063)**	0.051 (0.051)
	JMR	**0.223 (0.052)**	**0.310 (0.036)**	0.071 (0.040)
	JAMS	0.078 (0.049)	**0.149 (0.032)**	−0.032 (0.037)
	JCR	**0.278 (0.056)**	**0.347 (0.048)**	**0.180 (0.041)**
Article effects	Sub_newprod	**0.165 (0.069)**	0.094 (0.070)	**0.158 (0.057)**
	Sub_b2b	**−0.130 (0.065)**	−0.060 (0.074)	**−0.166 (0.060)**
	Sub_relationship	**0.285 (0.066)**	**0.295 (0.041)**	**0.216 (0.052)**
	Sub_prodbrand	−0.034 (0.030)	−0.032 (0.037)	**−0.079 (0.032)**
	Sub_ad	−0.128 (0.061)	−0.073 (0.059)	**−0.207 (0.055)**
	Sub_pricing	−0.079 (0.056)	**−0.115 (0.059)**	**−0.203 (0.045)**
	Sub_promotion	−0.123 (0.068)	**−0.158 (0.065)**	−0.138 (0.071)
	Sub_retailing	−0.002 (0.048)	0.071 (0.050)	**−0.135 (0.040)**
	Sub_strategy	**0.113 (0.041)**	0.027 (0.045)	0.024 (0.035)
	Sub_sales	−0.169 (0.096)	**−0.175 (0.083)**	**−0.292 (0.082)**
	Sub_method	−0.048 (0.029)	**−0.067 (0.025)**	**−0.076 (0.025)**
	Sub_services	**0.264 (0.049)**	**0.303 (0.044)**	**0.116 (0.035)**
	Sub_cb	**0.095 (0.025)**	**0.118 (0.034)**	−0.018 (0.033)
	Sub_international	0.098 (0.078)	**0.158 (0.066)**	0.028 (0.076)
	Sub_ecom	**0.495 (0.105)**	**0.463 (0.090)**	**0.452 (0.076)**
Journal effect*author effect		**−0.066 (0.010)**	**−0.024 (0.011)**	**−0.085 (0.010)**
Variances in Tobit		**2.183 (0.083)**	**0.200 (0.023)**	**0.339 (0.030)**

Table 4 Estimation results for the starting probability equations

Variables	State 1	State 2
Intercept	18.134 (17.538)	**−28.615 (8.921)**
No of authors	**75.454 (7.799)**	**59.110 (6.646)**
Editbrdjm	**61.943 (9.033)**	**71.011 (23.265)**
Editbrdjmr	**57.288 (10.158)**	−0.923 (24.826)
Editbrdjcr	−29.026 (20.385)	14.206 (11.106)
Editbrdmks	10.091 (12.705)	9.108 (11.456)
JM	15.600 (11.672)	−3.225 (19.236)
IJRM	**71.181 (12.922)**	**−49.984 (15.815)**
MKS	**76.177 (14.181)**	**53.519 (6.920)**
JMR	**60.513 (10.108)**	**35.640 (13.313)**
JAMS	9.921 (9.628)	11.644 (7.598)
JCR	−7.315 (7.479)	**20.358 (9.701)**

second state. As expected, winning prestigious awards leads the article to stay for less time in the low state and jump to the high state, as well as stay longer in the second state due to higher perceived quality from the scholars.

The estimate for the log rank is significantly positive in State 1 but not in State 2, which shows that low-ranking articles (with high log rank value) tend to stay in the low state longer, possibly due to lower perceived quality. Consistent with our expectation, we find there is a significant Matthew effect for marketing articles in the data. Specifically, the posterior mean estimate for the coefficient of the square of log rank is significantly positive in State 1 while negative in State 2, indicating the longer duration in the high state versus low state for the reverse-scored high ranking papers. That is, when an article becomes more famous with higher rank (i.e., lower value of log rank), it results in higher perceived quality and hence stays in the high state longer, which in turn disproportionately attracts more excess citations because of fame alone.

The estimate for the long number of years with zero citation is significantly negative in State 1 but not in State 2. It indicates that zero-cited papers tend to stay shorter in the first state and jump to the high state with the passage of time. However, we find presence of significant uncitedness effect for the marketing articles we analyzed. The posterior mean estimate for the coefficient of the square of log number of years with zero cites is significantly positive in State 1 but not in State 2, indicating that as time passes by, after certain point, the zero-cited papers tend to stay in the low state longer and longer. That is, zero-cited papers are more and more likely to be perceived to be relatively low quality and hence may be less and less likely to get cited over time.

For comparison purposes, we also include the main estimation results of the one-state discretized Tobit model in Table 3, which is the same as the heterogeneous discretized Tobit model without accounting for the latent citability states. It is clear from Table 3 that there exist significant estimation biases in the one-state model compared to the two-state model, which may lead to incorrect managerial implications (see Appendix C).

Dynamics of HMM states

Next, we examine how the HMM states dynamically change over time and how the predicted and actual citations differ by the HMM states over time. Figure 2 shows the estimated average probability of an article being in the high citability state—State 2 based on the estimation sample as well as the logarithm of actual new citations over time. It is interesting to see that as the actual new citations keep increasing over time, an article typically has a low probability to start in the high state (about 0.22) in the first two years, then jumps to 0.60 in the third year and stays relatively stable afterwards. This indicates the importance for authors of getting the article into the high citability state early on.

Figure 3 shows the actual and predicted new citations by the HMM states over time. It is clear from this figure that the citation dynamics are quite different across the two HMM states. In the high state, the annual new citations of articles tend to grow much faster than those in the low state. The predicted new citations in the two states seem to match the actual citations quite well. Again, this demonstrates the importance of both across and within article heterogeneity (i.e., the dynamics of HMM states) in explaining the citation dynamics of articles.

Based on the estimation and predication results, we next outline the theoretical contributions and implications. We focus on the interesting insights on dynamic citation prediction and the article citation dynamics which are not examined in the marketing literature.

Table 5 Estimation results for the waiting time equations

Variables	State 1	State 2
Intercept	**−21.695 (11.446)**	−14.339 (8.665)
Time	**−35.468 (7.146)**	**−45.446 (10.256)**
Square of Time	**18.705 (3.989)**	−14.069 (10.001)
No of Awards	**−43.827 (4.448)**	**22.703 (9.270)**
Log(rank)	**36.522 (13.552)**	9.984 (18.690)
Log^2(rank)	**74.763 (6.020)**	**−42.328 (7.182)**
Log(no of years with zero cites)	**−56.313 (14.742)**	4.218 (7.240)
Log^2(no of years with zero cites)	**56.715 (8.784)**	24.317 (18.742)

Summary

As Clark et al. (2013) and Stremersch et al. (2007) point out, marketing scholars need to pay more attention to article impact since it is in essence a “marketing of science” problem. Citations are an important measure widely used to assess the quality and influence of an article. Although there exists much research on citation measures and drivers in the marketing and scientometric literature (Stremersch et al. 2007), to the best of our knowledge, there is little recognition or research on the dynamic nature of citation influence. Our model allows institutions, scholars, journals, and scientometric scholars to understand the citation dynamics and estimate the future impact of a published article. The implications of our model are multidisciplinary in that the model can be used by scholars in varied disciplines.

Theoretical contributions

Relying on the literature on relationship marketing, we introduce to the literature the construct of article citability and suggest that article influence is determined by its relational embededness and the relational state an article is in. We empirically demonstrate the influence exerted by this latent factor on eventual article influence. The empirical results show that the discretized Tobit model with two-state HMM process outperforms all other competing models including the widely used NBD models in our context. We find that an article typically starts (with high probability) in the relatively low citability state in the publication year. On average, both latent states have high persistence such that an article will stay in either state for more than five years before jumping to another state, with slightly longer duration in the high state. Different drivers of citations exert differential influence on citability in the two states, and these influence dynamics change over time. Article citability moderates the influence of various known effects on final citation count.

Most models of citations such as the NBD model assume that the influence drivers are static and do not change over time with changes in an article’s latent state. We demonstrate

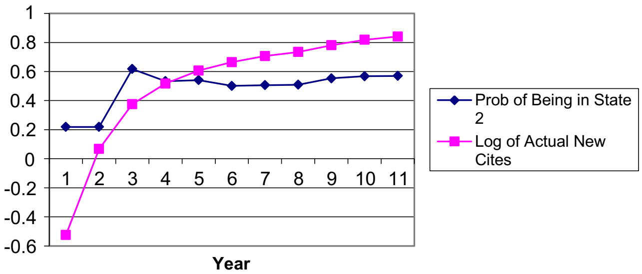


Fig. 2 Dynamics of HMM states over time

the shortcoming of this assumption and identify why it is critical to capture these dynamic changes. We recommend that researchers use our dynamic HMM model as it outperforms the NBD model and more accurately reflects citation dynamics.

Our study not only confirms that the trend of increasing yearly citations to the highly cited articles suggests that citations beget more citations, it but also provides a behavioral explanation for why the article-level Matthew effect exists. That is, highly cited articles result in a higher citability state and hence lead to more citations. This is different from the disproportionate attention rationale proposed by Small (2004) which suggests first citations reflect expert judgments of the contribution of an article: "The operation of this feedback mechanism, once set into motion will increase the inequality of citations by focusing attention on a smaller number of selected sources, and widening the gap between symbolically rich and poor" (p. 74). Our findings suggest that the article-level Matthew effect has a positive effect that improves latent citability states of the article. More importantly, we demonstrate that the Matthew effect is not merely a "famous author" effect but there is a "famous paper" effect as well. Our model outperforms the commonly used NBD model and can account for zero citations explicitly. Furthermore, to the best of our knowledge, no study has investigated the existence and dynamics of an uncitedness effect and the article-level Matthew effect on the citation dynamics of articles. To fill this knowledge gap, we proposed a conceptual framework of article citability and adopted a modified version of the model proposed by Li et al. (2005) to model individual article citations over time. Based on the citation literature, we

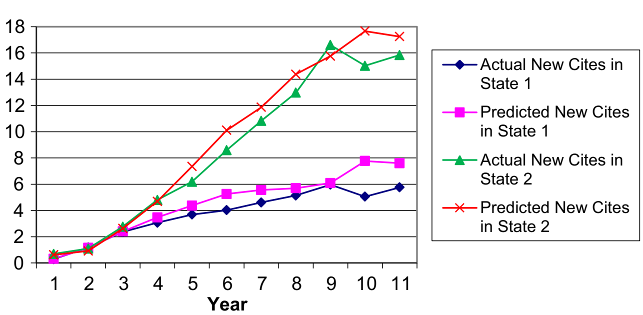


Fig. 3 Actual and predicted new cites by HMM states

propose appropriate measures to quantify both the article-level Matthew effect and the uncitedness effect.

Zinkhan (2005) suggests zero citation is a pattern characteristic of the marketing literature: "Over time, a few articles emerge as influential and frequently cited and the majority of articles sink into obscurity" (p. 253). However, zero citation is not unique to marketing. Garfield (2005) estimates that approximately 50% of the articles in the entire ISI database have zero citations. The NBD model and other models that purport to uncover the dynamics of citation fail to account for the impact of uncitedness. We view this as a critical limitation of those models.

Implications

All seven of the journals in our sample are highly regarded journals. Our findings suggest that articles start out in one of two citability states and these two states are not simply a case of elite (A-level) journals starting out in higher citability states in comparison to the near-elite (A-minus) journals. Citability states and article influence are not merely about the ranking of journals.

We find that quantitatively oriented articles (those published in *MKS*, *JMR*, and *IJRM*) tend to start and stay in the lower citability state. One important implication for marketing journals (particularly *MKS*, *JMR*, *IJRM*) will be examining why articles published at these outlets are in the relatively low citability state in the year of the article publication compared to the base journal—*JR*. Or, to state this differently, *JR* articles have a higher citability than the more highly regarded *MKS* and *JMR*. We conjecture that the more quantitatively oriented *MKS* and *JMR* may have more articles using sophisticated techniques that fewer researchers fully comprehend. Clark et al. (2013) comment that the declining influence of marketing scholarship could be attributed in part to "methodological sophistication" taking precedence over substantive issues. Editors at *MKS* and *JMR* should pay closer attention to the substantive and theoretical aspects of papers being submitted to their journals.

To further explore this we did some post-hoc analysis on *JR*. While *JR* is not positioned as a highly quantitative/modeling journal per se, it does publish modeling papers frequently. So we compared the citations accrued to modeling versus all other papers at *JR*. A t-test indicated that modeling papers in *JR* were less influential than non-modeling papers published there. This leads us to conclude that journals such as *JMR* and *MKS* that focus on highly sophisticated techniques should explore ways to make their findings more accessible by publishing an expanded non-technical summary of the findings and paying closer attention to the theoretical contributions of work submitted there. The greater citability state and higher probability of starting in the higher state of *JCR* compared to all other leading marketing journals is something that needs to be explored further. *JCR* is positioned as an interdisciplinary journal and may perhaps draw attention from a broader array of researchers. The greater influence dynamics of *JAMS* relative to *MKS* and *JMR* is also worth noting.

Further, based on the estimated main effects of the journals on new citations, it is important to note that only *JM* and *JCR* have significant impact in both high and low citability states, while *MKS*, *JMR*, and *JAMS* increase citations in the low state only, and *IJRM* decreases citations in both states compared to *JR*. Therefore, the latter four journals may again want to boost their relative citability states in scholars' minds. *IJRM* is accorded an A-minus status in many institutions but seems to have considerably less citability than its peer journals.

We find the presence of a strong article-level Matthew effect and uncitedness effect for marketing articles in that they increase or decrease the article's citability, which leads to more or less influence for the article, respectively. There are significant author effects, journal effects, and article effects on articles' influence dynamics, and these effects dynamically change over time according to the latent citability state in which the article is. The effect of author reputation is muted or considerably diluted when the article is in a low citability state. Also, for papers published in A-level marketing journals or for papers starting out in the high citability state, author editorial board membership and author fame are less critical for eventual influence. This indicates that author fame is less important when a paper is published in an A-level marketing journal, especially in the high citability state. Further, we find that editorial board membership at *JM* or publishing in *JCR* results in a higher citability state. And articles published in *JM* or *JCR* seems to be in relatively higher citability states compared to other leading marketing journals. Contrary to our expectations, enhanced citability state of an article mitigates the journal effects. When an article is in a low citability state, it benefits from stronger journal effects compared to being in high citability state. When that same article moves into a high citability state, the journal effects become weaker possibly due to some ceiling effect.

We find that the number of authors significantly decreases the number of new citations in the high citability state while it has insignificant impact in the low citability state. We also find in our context that an article typically has a low probability of starting in the high state (about 0.22) in the first 2 years, then jumps to 0.60 in the third year and stays relatively stable afterwards. In the high state, the annual new citations of articles tend to grow much faster than those in the low state. This indicates the importance for authors of getting the article into the high citability state early on. Our research indicates that if a paper is not cited early, while it may not go uncited in the future, it may result in low relational state in scholars' minds, and hence its long-term influence is in great doubt. Therefore, our study has important managerial implications for journals, institutions, and scholars.

Limitations

We note the relative newness of Scopus in comparison to the ISI database. There is a large literature that has investigated problems in the ISI database such as the numerator/denominator discrepancy, but Scopus was launched in 2004 so there is no research indicating what potential biases and anomalies may be present in this database. There is also a problem of an absence of comparable Scopus data; previous studies using ISI data such as Zinkhan and Leigh (1999) provide statistics from earlier years that can be used for comparative purposes. Another difficulty is that Scopus citations for social science journals have been extended back only to 1996, which is insufficient for determining the time in which citations to these marketing articles may decline and exhibit ageing or obsolescence.

Another limitation of this study is the focus on a small group of seven core marketing journals. Based on an analysis of economic journals, Davis (1998) argues there are substantial differences in citation impact patterns for general versus specialized journals within a discipline. Further research is needed to investigate patterns characteristic of specialized marketing journals such as the *Journal of Consumer Psychology*, *Journal of Advertising*, or *Industrial Marketing Management.* Davis (1998) also argues for citation impact differences between disciplinary and interdisciplinary journals. Interdisciplinary business journals such as *Harvard Business Review*, *Sloan Management Review*, and *Management Science* may have different citation impact characteristics than the marketing-focused journals. While we did analyze 1,591 articles published in seven journals, we acknowledge that these articles probably represent around 15–20% of all articles published in marketing journals during this time frame.[2]

The article ranking used to capture the article-level Matthew effect in this study is based on the sample. The Matthew effect may be even stronger with larger sample with more articles or journals included. Further, due to data unavailability, we do not observe self-citation information for the articles in this study. Future research could investigate the dynamics of self-citations and the difference from non-self-citations over time. While we find that the number of co-authors is a net positive for eventual influence, it does raise interesting questions regarding the relative contributions of individual authors. Sahu and Panda (2014) suggest that more co-authors can foster interdisciplinary work and boost the overall quality of the paper. But on the other hand, Persson and Glanzel (2014) raise the specter of honorific co-authoring

[2] Hult et al. (1997) ranked 29 marketing journals, ten business journals, and two proceedings. The 7 journals we studied were among the 39 journals that were ranked. We conservatively estimate that the number of articles analyzed would be about 15–20% of articles published during 1996–2003 in the 39 journals ranked by Hult et al.

that might emerge when authorship is credited to those making the most minimal of contributions. We commend researchers to investigate this.

While citation analysis has been viewed as superior and more objective than simply collating opinions, it has its limitations and this needs to be acknowledged. Sometimes citations may not measure intellectual influence but may reflect criticisms of a paper or temporary faddish interest in certain topics (Hofacker et al. 2009). Our paper focused on aggregate scholar/fields citation practices and did not focus on individual reader behaviors (Hofacker et al. 2009). Finally, our focus was on the impact of academic research on the literature. In applied disciplines, scholarship can also have impact on practice (Reibstein et al. 2009; Schultz 2012). Citation based analysis does not measure this type of impact.

Appendix A

Journal differences and citation dynamics

Table 6 presents the citation measure summary for the seven journals in the data. Clearly, there exist big differences across the journals. *JM*, *JMR*, and *JCR* have the highest mean or median number of citations, highest percentage of articles with 20 or more cites, and lowest percentage of zero-cited articles, followed by *JR* and *JAMS*, then *MKS* and *IJRM*.

To see if different articles have different citation dynamics, we divide the articles into two groups based on the average total number of citations received at the end of the data period: articles with higher than average total citations versus those with lower. Then we plot the annual average new citations over time for the two groups in Fig. 4 and the annual change of average new citations by year in Fig. 5. From Fig. 4, it is clear that the citation dynamics for these two sets of articles are dramatically different. Those articles with high total citations in the end tend to grow much faster over time than those with low total citations. For the latter, it seems that their new citations initially increase and quickly become somewhat stable after 4 or 5 years of publications with even slightly declining trend after that. Figure 5 further confirms the different citation dynamics for the two sets of articles. Also, we can clearly see from Fig. 5 that the within-article citation change varies dynamically over time. For instance, for the articles with high total citations, the annual change of citations first increases in the first 3 years, then decreases afterwards with some bounce back in year 10. All these indicate that it is important to account for article heterogeneity and the citation dynamics both within and across articles.

Appendix B

Likelihood and identification of the proposed model

The unconditional likelihood is given by (Liechty et al. 2003):

$$\mathrm{Prob}(\mathbf{Z}) = \prod_i \prod_s (\pi_i(s))^{\mathrm{I}\{D_{i0}=s\}} \left\{ \prod_t \mathrm{Prob}(Z_{it}|\mathrm{D}_{it}) \cdot \left(\prod_l P_{itsl}{}^{\mathrm{I}\{D_{it}=s \ \& \ D_{i(t+1)}=l\}} \right) \cdot \prod_k f(W_{itk}|D_{it}, \lambda_{it}, \kappa_i) \right\},$$

where $\mathbf{Z}$ is the whole observed citation sequence across articles and time, π_i is the starting probability at time 0, and P_{itsl} is the transition probability from state s to l at time t defined in the HMM process. $I\{\}$ is an indicator function. D_{it} is the realized latent citability state for article i at time t. $f(W_{itk}|D_{it}, \lambda_{it}, \kappa_i)$ is the density of the k-th waiting time of article i at time t conditional on the realized hidden state D_{it} and Gamma shape and scale parameters λ_{it} and κ_i, which is given by

$$f(W_{itk}|D_{it}, \lambda_{it}, \kappa_i) = \zeta_k \sum_{m=0}^{\infty} \frac{\kappa_i{}^{\lambda_{ik}(D_{it})}}{\Gamma(\lambda_{ik}(D_{it}))} (W_{itk}(D_{it}) - \tau_{k+m})^{\lambda_{ik}(D_{it})-1} \exp\{-\kappa_i(W_{itk}(D_{it}) - \tau_{k+m})\} \cdot \mathrm{I}\{\tau_{k+m} < W_{itk}(D_{it}) \le \tau_{k+m+1}\}.$$

Table 6 Citation difference across journals

Influence measure	JM	JMR	JCR	MKS	JAMS	JR	IJRM
# of articles 96-03	241	302	282	189	212	180	185
Mean number of Citations	32.25	17.63	16.03	11.45	15.66	16.40	8.46
Median number of citations	19.0	12.0	11.0	7.0	10.0	10.0	5.0
Maximum number of cites	423	108	149	68	139	110	94
Percent uncited	1.6	1.0	2.8	4.4	1.9	3.6	5.8
Percent with 20 or more cites	49.0	29.5	27.8	15.8	23.3	25.5	9.0

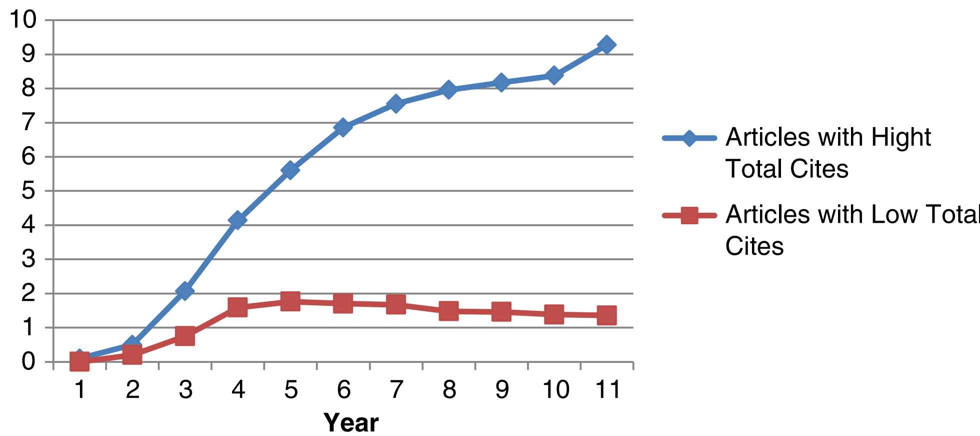


Fig. 4 Average new cites by year

Here $\lambda_{ik}(D_{it})$ is the Gamma shape parameter associated with the state of an article when the k-th jump time begins and with the realized state D at time t, and τ_{k+m} is the m-th time after τ_k—the k-th jump time of the article—that D changed states, with $\tau_{k+0}=0$. ζ_k is a normalizing constant of the density.

To ensure identification of the hidden states, we restrict the average new citations to be non-decreasing in the states (Li et al. 2011; Netzer et al. 2008). That is, we impose this restriction at the intercepts of Eq. 2 in the text such that $\gamma_{0i1}\leq\gamma_{0i2}\leq\cdots\leq\gamma_{0iS}$. We also refer scholars interested in the estimation procedure to Li et al. (2005), Liechty et al. (2003) and Li et al. (2011).

Appendix C

Comparison to the one-state model results

For comparison purposes, we include the main estimation results of the one-state discretized Tobit model in Table 3 in the text. Note that the one-state model is the same as the heterogeneous discretized Tobit model without accounting for the citability states. It is clear from Table 3 that there exist significant estimation biases in the one-state model compared to the two-state model, which may lead to incorrect managerial implications. For instance, without incorporating the dynamics of citability states of articles, the one-state model tends to over-estimate the persistence effect (p-value = 0). Regarding the author effects, based on the one-state model, authors may incorrectly conclude that the number of co-authors does not matter in attracting new citations but it actually may hurt when the article is in the high citability state. The one-state model also tends to over-estimate the impact of the editorial board membership at *JM*, *JMR*, *JCR*, and *MS* compared to those in the low state in the two-state model while under-estimating their impacts compared to those in the high state. For the journal effects, there exist significant under-estimation biases when compared to those in the low state but over-estimation biases when compared to those in the high state in the two-state model. Similarly, for the article effects, significant estimation biases are also present in the one-state model. For example, the topics of new products and business to business marketing are shown to have significant impacts on new citations in the one-state model while their effects are only significant in the high state in the two-state model. Many subject areas (i.e., product and brand management, advertising, pricing, promotion, retailing, sales, methodology, international marketing) are shown to have insignificant impact in the one-state model but have significant impact

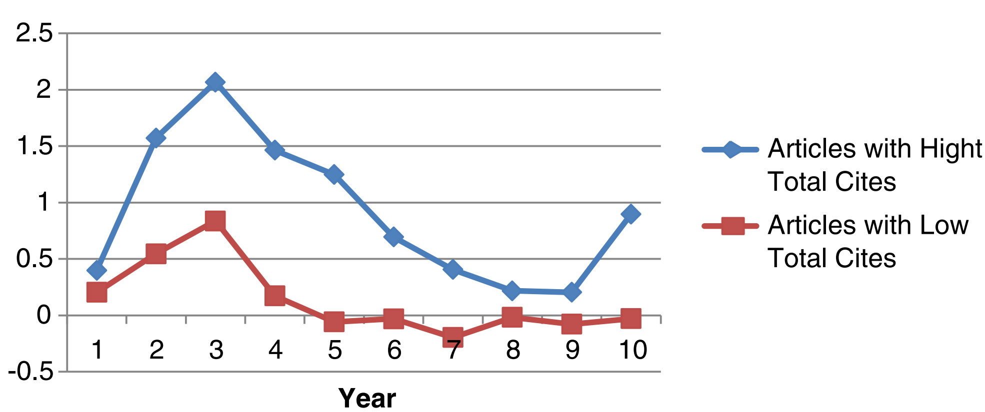


Fig. 5 Annual change of average new cites by year

Table 7 Concentration of citations to the top five articles by journal and year of publication

Year of article publication	JM	JMR	JCR	MKS	JAMS	JR	IJRM
Citations to top five of 1996	1016	269	204	316	227	192	299
% of total citations of 1996 articles	44.1%	35.3%	28.6%	52.9%	56.0%	53.8%	57.6%
Citations to top five of 1997	794	462	279	218	440	233	135
% of total citations of 1997 articles	54.5%	40.7%	27.8%	42.1%	68.6%	49.8%	49.3%
Citations to top five of 1998	554	212	549	270	163	187	126
% of total citations of 1998 articles	34.6%	24.9%	47.3%	56.5%	50.3%	62.8%	51.9%
Citations to top five of 1999	448	336	200	212	258	94	59
% of total citations of 1999 articles	29.1%	34.8%	33.7%	39.9%	42.3%	49.0%	54.6%
Citations to top five of 2000	299	191	161	346	237	203	131
% of total citations of 2000 articles	33.6%	35.4%	24.2%	72.2%	32.5%	59.2%	66.2%
Citations to top five of 2001	192	227	162	113	119	119	55
% of total citations of 2001 articles	32.6%	38.0%	26.5%	51.6%	57.5%	46.7%	56.7%
Citations to top five of 2002	157	98	104	81	104	68	58
% of total citations of 2002 articles	28.6%	34.6%	24.2%	39.3%	30.8%	58.6%	61.1%
Citations to top five of 2003	91	78	69	65	50	36	41
% of total citations of 2003 articles	32.6%	40.4%	28.9%	34.9%	45.0%	54.5%	64.1%
Mean concentration of citations to top five 1996-2003	36.2%	35.5%	30.2%	48.7%	47.9%	54.3%	57.7%

All the citation numbers in the table are as of 2006

either in low state or high state in the two-state model. Lastly, the one-state model seems to over-estimate the variance of the new citations compared to those in the two-state model. These estimation biases in the one-state model highlight the importance of accounting for the dynamics of latent citability states of articles over time.

Appendix D

Concentration of citations to the top five articles by journal

The effects of "hit" papers are also indicated in statistics on concentration of citations to the top five articles of the year. Table 7 presents an overview of the extent of concentration by year for these selected journals. *MKS* in 2000 records the highest concentration of citations (72.2%), reflecting the success of a special issue on marketing science and the Internet. Concentration in this case is the percentage of citations for the five most cited articles of the year relative to total citations for all articles in the year. The results suggest the journals with lower prestige may be consistently high in their concentration percentages or dependencies on their most cited articles. For example, concentration for *IJRM* ranges from 49.3 to 66.2% over the period of 1996–2003.

There is also significant variability year-to-year reflecting the publishing of "hit" articles or the lack of "hits" in a year. For example, the top five 1999 articles in *MKS* account for only 39.9% of cites. However, in 2000 the concentration jumps to 72.2%, driven by the success of three articles on internet shopping from a special issue of *MKS* (Novak et al. 2000; Lynch and Ariely 2000 and Haubl and Trifts 2000). *JCR* has the lowest average concentration of these seven journals. However, *JCR* concentration jumps from 27.8% in 1997 to 47.3% in 1998. The year 1998 includes the only three *JCR* articles with more than a hundred citations in this dataset (Bettman et al. 1998; Fournier 1998 and Steenkamp and Baumgartner 1998).

References

Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the web of science and scopus. *Journal of the American Society for Information Science and Technology, 60*(7), 1320–1326.

Baldi, S. (1998). Normative versus social constructivist process in the allocation of citations: a network-analytic model. *American Sociological Review, 63*(December), 829–846.

Baumgartner, H., & Pieters, R. (2003). The structural influence of marketing journals: a citation analysis of the discipline and its sub-areas over time. *Journal of Marketing, 67*, 123–139.

Bergh, D. D., Perry, J., & Hanke, R. (2006). Some predictors of SMJ article impact. *Strategic Management Journal, 27*, 81–100.

Berry, L. L. (1983). Relationship marketing. In L. Berry, G. L. Shostack, & G. D. Upah (Eds.), *Emerging perspectives on services marketing* (pp. 25–58). Chicago: American Marketing Association.

Bettencourt, L. A., & Houston, M. B. (2001). The impact of article method type and subject area on article citations and reference diversity in *JM, JMR,* and *JCR*. *Marketing Letters, 12*(4), 327–340.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice process. *Journal of Consumer Research, 25*(December), 187–217.

Bonitz, M., Bruckner, E., & Scharnhorst, A. (1997). Characteristics and impact of the Matthew effect for countries. *Scientometrics, 40*, 407–422.
Burrell, Q. L. (2003). Predicting future citation behavior. *Journal of the American Society for Information Science and Technology, 54*(5), 372–378.
Chib, S. (1995). Marginal likelihood from the Gibbs output. *Journal of the American Statistical Association, 90*(432), 1313–1321.
Chib, S., & Greenberg, E. (1995). Understanding the Metropolis Hastings algorithm. *American Statistician, 49*, 327–335.
Chow, C. W., Haddad, K., Singh, G., & Anne, W. (2007). On using journal rank to proxy for an article's contribution or value. *Issues in Accounting Education, 22*(3), 411–427.
Clark, T., Key, T., Hodis, M., & Rajaratnam, D. (2013). The intellectual ecology of mainstream marketing research: an inquiry into the place of marketing in the family of business disciplines. *Journal of the Academy of Marketing Science*. doi:10.1007/s11747-013-0362-5.
Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago: University of Chicago Press.
Cronin, B. (1981). The need for a theory of citing. *Journal of Documentation, 37*, 16–24.
Damien, P. S., Wakefield, J., & Walker, G. (1999). Gibbs sampling for Bayesian nonconjugate and hierarchical models using auxilliary variables. *Journal of Royal Statistical Society, Series B, 61*(Part 2), 331–344.
Davis, J. B. (1998). Problems in using the *social sciences citation index* to rank economic journals. *American Economist, 42*(Fall), 59–65.
Du, R., & Kamakura, W. (2006). Household life cycles and lifestyles in the United States. *Journal of Marketing Research, 43*(1), 121–132.
Dwyer, F. R., Schurr, P. H., & Sejo, O. (1987). Developing buyer-seller relationships. *Journal of Marketing, 51*(April), 11–27.
Fournier, S. (1998). Consumers and their brands: developing relationship theory in consumer research. *Journal of Consumer Research, 24*(March), 343–373.
Garfield, E. (1979). *Citation indexing—its theory and application in science, technology, and humanities*. New York: John Wiley & Sons.
Garfield, E. (2005). The agony and the ecstasy—the history and the meaning of the Journal Impact Factor. http://garfield.library.upenn.edu/papers/jifchicago2005.pdf.
Gelfand, A., & Smith, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*, 398–409.
Guidry, J. A., Guidry, B. N., Holler, L. J., Tanner, J. R., & Veltsos, C. (2004). Surveying the cites: a ranking of marketing journals using citation analysis. *Marketing Education Review, 14*(Spring), 45–59.
Haubl, G., & Trifts, V. (2000). Consumer decision making in online shopping environments: the effects of interactive decision aids. *Marketing Science, 19*(Winter), 4–21.
Hendrix, D. (2008). An analysis of bibliometric indicators, National Institutes of Health Funding, and faculty size at Association of American Medical Colleges, medical schools. *Journal of Medical Library Association, 96*(4), 324–334.
Hult, G., Tomas, M., Neese, W. T., & Edward Bashaw, R. (1997). Faculty perceptions of marketing journals. *Journal of Marketing Education, 19*(1), 37–52.
Hofacker, C., Gleim, M. R., & Lawson, S. (2009). Revealed reader preference for marketing journals. *Journal of the Academy of Marketing Science, 37*(2), 238–247.
Ijiri, Y., & Simon, H. A. (1974). Interpretations of departures from the Pareto curve firm-size distributions. *Journal of Political Economy, 82*, 315–331.
Katerratanakul, P., Han, B., & Hong, S. (2003). Obective quality ranking of computing journals. *Communications of the ACM, 46*(October), 111–114.
Keller, K. L. (2012). Understanding the richness of brand relationships: research dialogue on brand as intentional agents. *Journal of Consumer Psychology, 22*, 186–190.
Lehmann, D. R. (2005). Journal evolution and the development of marketing. *Journal of Public Policy & Marketing, 24*(1), 137–142.
Li, S., Liechty, J. C., & Montgomery, A. L. (2005). *Modeling category viewership of web users with multivariate count models, working paper*. Bloomington: Kelley School of Business, Indiana University.
Li, S., Sun, B., & Montgomery, A. L. (2011). Cross-selling the right product to the right customer at the right time. *Journal of Marketing Research, 48*(4), 683–700.
Liechty, J., Pieters, R., & Wedel, M. (2003). Global and local covert visual attention: evidence from a bayesian hidden Markov model. *Psychometrika, 68*(4), 519–541.
Luo, A., & Kumar, V. (2013). Recovering hidden buyer-seller relationship states to measure the return on marketing investment in business-to-business marketing. *Journal of Marketing Research, Vol. L*, 143–160.
Lynch, J. G., Jr., & Ariely, D. (2000). Wine online: search costs affect competition on price, quality, and distribution. *Marketing Science, 19*(Winter), 83–103.
Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using Clickstream Dat'. *Marketing Science, 23*(4), 579–595.
Moon, S., Kamakura, W. A., & Ledolter, J. (2007). Estimating promotion response when competitive promotions are unobservable. *Journal of Marketing Research, 44*(3), 503–515.
Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing, 58*(July), 20–38.
Medoff, M. H. (2006). Evidence of a Harvard and Chicago Matthew effect. *Journal of Economic Methodology, 13*(4), 485–506.
Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(January 5), 56–63.
Mingers, J., & Burrell, Q. L. (2006). Modeling citation behavior in management science journals. *Information Processing and Management, 42*, 1451–1464.
Mort, G. S., McColl-Kennedy, J. R., Kiel, G., & Soutar, G. N. (2004). Perceptions of marketing journals by senior academics in Australia and New Zealand. *Australasian Marketing Journal, 12*(2), 51–61.
Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing Science, 27*(2), 185–204.
Novak, T. P., Hoffman, D. L., & Yung, Y.-F. (2000). Measuring the customer experience in online environments: a structural modeling approach. *Marketing Science, 19*(Winter), 22–42.
Persson, O., & Glanzel, W. (2014). Discouraging honorific authorship. *Scientometrics, 98*(2), 1417–1419.
Pieters, R., & Baumgartner, H. (2002). Who talks to whom? Intra- and interdisciplinary communication of economic journals. *Journal of Economic Literature, 40*(2), 483–509.
Reibstein, D. J., Day, G., & Wind, J. (2009). Guest editorial: is marketing academia losing its way? *Journal of Marketing, 73*(4), 1–3.
Rossi, P. E., McCulloch, R. E., & Allenby, G. M. (1996). The value of purchase history data in target marketing. *Marketing Science, 15*(4), 321–340.
Sahu, S. R., & Panda, K. C. (2014). Does the multi-authorship trend influence the quality of an article. *Scientometrics, 98*(3), 2161–2168.
Samiee, S., & Chabowski, B. (2012). Knowledge structure in international marketing: a multi-method bibliometric analysis. *Journal of the Academy of Marketing Science, 40*(2), 364–386.
Schultz, D. E. (2012). The academic silly season. *Marketing News*, September 30, 14.

Seetharaman, P. B. (2004). Modeling multiple sources of state dependence in random utility models: a distributed lag approach. *Marketing Science, 23*(2), 263–271.

Seggie, S. H., & Griffith, D. A. (2009). What does it take to get promoted in marketing academia? Understanding exceptional publication productivity in the leading marketing journals. *Journal of Marketing, 73*(January), 122–132.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*(9), 628–638.

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *British Medical Journal, 314*(February 15), 498–502.

Simon, H. A. (1955). On a class of skew distribution functions. *Biometrika, 42*, 425–440.

Sivadas, E., & Johnson, M. S. (2005). Knowledge flows in marketing: an analysis of journal article references and citations. *Marketing Theory, 5*(4), 339–361.

Small, H. (2004). On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics, 60*(1), 71–79.

Spence, M. A. (1973). Job market signaling. *Quarterly Journal of Economics, 87*, 355–374.

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(June), 78–90.

Stern, R. E. (1990). Uncitedness in the biomedical literature. *Journal of the American Society for Information Science, 41*(3), 193–196.

Stremersch, S., Verniers, I., & Verhoef, P. C. (2007). The quest for citations: drivers of article impact. *Journal of Marketing, 71*(July), 171–193.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association, 82*(398), 528–540.

Tol, R. S. J. (2009). The Matthew effect defined and tested for the 100 most prolific economists. *Journal of the American Society for Information Science and Technology, 60*(2), 420–426.

Varadarajan, P. R. (2003). From the editor: Journal of the Academy of Marketing Science, 2 2000 to 2003. *Journal of the Academy of Marketing Science, 31*(4), 365–367.

Weale, A. R., Bailey, W., & Lear, P. A. (2004). The level of non-citation of articles within a journal as a measure of quality: a comparison to the impact factor. *BMC Medical Research Methodology, 4*, 14.

Woodside, A. G. (2009). Journal and author impact metrics: an editorial. *Journal of Business Research, 62*(1), 1–4.

Zinkhan, G. M. (2005). From the editor: scientific status and knowledge use: two perspectives. *Journal of the Academy of Marketing Science, 33*(Summer), 251–253.

Zinkhan, G. M., & Leigh, T. W. (1999). Assessing the quality ranking of the *Journal of Advertising. Journal of Advertising, 28*(Summer), 51–70.