Debates and Assumptions about Motion Picture Performance: A Meta-Analysis

François A. Carrillat*

Associate Professor of Marketing

University of Technology Sydney

14-28 Ultimo Road

Ultimo, NSW 2007

Australia

Tel: +61 2 9514 3530

Francois.carrillat@uts.edu.au

Renaud Legoux

Associate Professor of Marketing

HEC Montréal

3000, Chemin de la Cote-Sainte-Catherine

Montréal, QC H3T 2A7

Canada

Tel: +1 514 340 6997

Renaud.legoux@hec.ca

Allègre Hadida

University Senior Lecturer in Strategy

Fellow of Magdalene College

University of Cambridge

Trumpington Street

Cambridge CB2 1AG

UK

Tel: +44(0)1223 339612

a.hadida@jbs.cam.ac.uk

*: contact author

Debates and Assumptions about Motion Picture Performance: A Meta-Analysis

Abstract

Across the many studies of motion picture box office success, unresolved debates and untested assumptions about the contributing factors persist. Using an accessibility–diagnosticity framework and a meta-analysis of 634 effect sizes from 150 studies, the current article seeks to clarify the relationships of star brand equity and product reviews (from consumers and critics) with box office success. The popularity of stars (market and media appeals) exerts a stronger impact on box office success than their artistic recognition (as per award nominations and wins) at the moment of a movie's release but not over its extended theatrical run. Whereas the impact of popular stars on box office success decreases over time, the influence of artistically recognized stars remains steady. The findings also identify a dual role for critics, who influence consumers' movie choice and predict box office performance by merely reflecting moviegoers' tastes. Finally, this study refutes the assumption that the impact of users' reviews strengthens over time, relative to critics' reviews.

Keywords Cinema, Meta-analysis, Star brand equity, Product reviews, Critics' reviews, Users' reviews, Cue diagnosticity

In modern economies, cinema box office (BO) revenues continue to be substantial and are expected to reach US$45.9 billion worldwide by 2018 (PwC 2015). Yet predictions of BO success remain challenging, because movies are experiential products whose quality cannot be assessed in advance (Basuroy et al. 2006; Eliashberg and Sawhney 1994; Joshi and Mao 2012). Several factors, such as the presence of movie stars or positive reviews, can reduce consumers' uncertainty (e.g., Akdeniz and Talay 2013; Elberse 2007). The accessibility and diagnosticity of these factors have led to their prominence in motion picture performance research, attracting considerable attention from scholars across academic disciplines (Joshi and Mao 2012; Litman 1983; Liu 2006; Lynch et al. 1988; Purohit and Srivastava 2001; Ravid 1999; Zuckerman et al. 2003).

The many studies on stars and product reviews[1] offer little established knowledge about actual impacts on motion picture performance. In particular, questions remain about the relative impact of star popularity versus artistic recognition, as well as whether professional movie critics influence moviegoers' choices or simply predict them (Basuroy et al. 2003; Eliashberg and Sawhney 1994). Furthermore, some widespread assumptions about this industry have not been tested empirically, including the popular sense that users' reviews are gaining more influence over BO performance relative to professional critics' reviews (Flanagin and Metzger 2013; Liu 2006) or that movie stars' aura has dimmed in recent years (Marich 2014; Merry 2015; *The Economist* 2016).[2] We seek to test these and similar assumptions, spanning the salient debates that we list in Figure 1 (numbered from 1–5 to facilitate discussions throughout this text).

[Insert Figure 1 about here]

Specifically, to address the ongoing debates and challenge some persistent assumptions, we conduct a meta-analysis of 634 correlation effect sizes that have appeared in 150 published and unpublished manuscripts over the past four decades. The literature sample spans marketing, sociology, cultural economics, strategy, organizational theory, and psychology. With this broad scale of investigation, we can leverage study heterogeneity and thereby examine, for example, different approaches to star brand equity, based on either the stars' popularity (i.e., market dimension measured by previous BO revenue or media dimension measured by inclusion in "influential stars" lists) or their artistic recognition (i.e., artistic dimension measured by peer-

---

[1] Reviews are often referred to as third-party evaluations (e.g., Basuroy et al. 2006; Chang and Ki 2005; Chen et al. 2012), but we opt for the more direct "product reviews" terminology for simplicity.
[2] Although Hofmann et al. (2017) consider this assumption, they obtain different results than we do, as we detail subsequently.

based awards). The vast and diverse pool of primary studies also enables us to determine if the relationship between BO revenue and product reviews depends on the review source (consumers or critics), metric (valence, or average evaluations, versus volume, or the number of reviews), or performance indicator (opening weekend, total BO, or other). With these considerations, our study builds on and extends a previous meta-analysis of star brand equity (Hofmann et al. 2017), as well as more general meta-analyses of product reviews (Babić Rosario et al. 2016; Floyd et al. 2014; You et al. 2015). Our findings conflict with some prior conclusions, reflecting the smaller pool of effect sizes and fewer statistical controls in previous studies. We highlight some key research differences in Table 1, which also underscores the three main contributions of this study.

[Insert Table 1 about here]

First, we distinguish between the popularity and artistic recognition of stars to determine their effects on both short- and long-term BO. Hofmann et al. (2017) assert that the market (popularity) dimension of star brand equity exerts a stronger impact on BO than the artistic dimension, but our results indicate that this difference holds only for short-term BO. In the longer run, the effects of both market and artistic dimensions are similar. Furthermore, we find a steady influence of stars' artistic recognition on theatrical success since the early 1930s, such that it represents the most stable dimension of stars' brand equity. In contrast, the impact of their popularity has been declining. The increasing variance of movie quality over time has seemingly offset this diminishing influence of star power (Johnston 2009), such that we can assert that casting stars in projects of similar quality, rather than attempting to salvage poor quality movies with powerful stars, helps maintain the diagnosticity of their popularity.

Second, we apply criteria proposed by Eliashberg and Shugan (1997) and Basuroy et al. (2003) to differentiate *influencer* and *predictor* role for critics, across review metrics. The results show that critics can both influence and predict BO revenue. Specifically, reviews exhibit a positive, significant association with both short- and long-term BO, whether we measure their volume or valence. Thus, studios should aim to have their movies reviewed by as many critics as possible. They also can rely systematically on the number of reviews to forecast their theatrical success.

Third, by comparing the effects of product reviews from different sources and with distinct metrics, we issue a strong rebuttal of the assumption that other consumers have more clout than critics. Whether we consider review volume or valence, consumers do not display stronger effects than critics, and this result remains stable over the years. Therefore, switching promotional efforts to target users at the expense of critics is unwarranted for studios, now or in the future, as we find no significant trends.

Conceptual background

The quality of experiential products such as movies is difficult to ascertain in advance, so consumers seek cues to diminish their uncertainty (Basuroy et al. 2006; Eliashberg and Sawhney 1994; Reddy et al. 1998). In particular, star brand equity stems from the popular appeal or artistic recognition of prominent actors, directors, composers, or producers (Albert 1998; Elberse 2007; Luo et al. 2010). Powerful stars with high brand equity function like branded attributes for movies (Hofmann et al. 2017; Levin et al. 1997; Liu et al. 2014), providing cues that help consumers evaluate a movie before they watch it (Hennig-Thurau et al. 2001) and heuristics for them to make their consumption choices (Desai and Basuroy 2005). In addition, consumers use product reviews to support their search for information about the quality of movies as experience goods (Nelson 1970). Product reviews by other moviegoers provide non-expert or amateur insights, mostly through posts online (e.g., IMDb.com) or social media (e.g., Twitter) (Basuroy et al. 2003; Hennig-Thurau et al. 2015). Reviews by critics instead represent expert, professional assessments, published in both off- and online media (Basuroy et al. 2003; Eliashberg and Shugan 1997).

Other constructs beyond star brand equity and product reviews can issue cues too, such as a film's budget, promotional expenses, genre, or plot. However, star brand equity and product reviews appear to offer the most salient cues (Desai and Basuroy 2005). Powerful stars are often instrumental in getting movie projects "green-lit" and exert substantial power when negotiating with studios (Anand and Watson 2004). They can establish immediate rapport with consumers by being both ordinary and extraordinary, allowing audiences to both identify with them and admire them, respectively (Sedgwick 1999). In turn, product reviews are the most prominent quality cues that are external to the production firm—unlike the stars, who are selected by the

studies (Hadida 2009). Expert reviews provide unique insights about the quality of experience goods (Basuroy et al. 2003); non-expert reviews represent trusted sources of information, such that 95% of consumers search for online reviews, and 84% trust them as much as personal recommendations (Brightlocal 2016).

Accessibility–diagnosticity

According to the accessibility–diagnosticity framework (Feldman and Lynch 1988; Lynch et al. 1988), consumers rely on the cues that are more diagnostic and accessible to make a judgment (Joshi and Mao 2012; Knapp et al. 2014; Langan et al. 2017). We use conceptual arguments based on this accessibility–diagnosticity framework to predict and establish which of the cues associated with star brand equity and product reviews are more likely to drive consumer choice and BO performance. We organize each discussion according to the prominent debates and assumptions surrounding each cue.

Star brand equity

Debate Which dimension of star brand equity has the strongest impact on theatrical performance (issue 1 in Figure 1)? One perspective, based on Stigler and Becker's (1977) theory of consumption capital, contends that the value of stars for moviegoers stems from their popularity, which might be defined by their market success or media prominence (Adler 1985). In this view, star brand equity derives from the utility that audiences obtain from being able to talk about familiar stars and their performances with friends or acquaintances in socially rewarding conversations that enhance their consumption capital. That is, popular stars represent a better source of consumption capital than less popular counterparts. Another perspective instead highlights artistic recognition, with the argument that consumers are drawn to movies that feature stars whose performance is enjoyable in and of itself, due to their artistic skills. Such stars often achieve artistic recognition, in the form of awards or nominations (Bagella and Becchetti 1999; Deuchert et al. 2005; Hofmann et al. 2017; Rosen 1981; Suárez-Vázquez 2011).

Empirical evidence is conflicted: Some studies find that the market dimension of star brand equity relates more strongly to BO than its artistic dimension (Liu et al. 2014), but others indicate that both forms matter (Hadida 2010). Popular stars often command higher fees, so it is important for studios to verify empirically whether they also lead to higher BO. Hofmann et al.'s (2017) meta-analysis, based on 172 effect sizes, indicates that the market dimension has a

stronger impact on BO revenue than the artistic dimension, but their investigation of this debate is incomplete. Namely, popularity consists of both market and media appeals (Deuchert et al. 2005), yet these authors assess market and artistic dimensions solely, without comparing the media and artistic dimensions.

Both popularity and artistic recognition seemingly should reduce quality uncertainty by providing diagnostic cues (Lynch et al. 1988). Popular stars enhance consumption capital and diminish the audience's risk that the movie will not deliver utility, and stars with artistic recognition indicate that the movie should be enjoyable (Suárez-Vázquez 2011). Popularity and artistic recognition also are equally accessible (Lynch et al. 1988) but offer different cues, reflecting the "commerce versus art" schism in creative industries (Holbrook and Addis 2007). Thus, they do not compete against each other but instead face competition with the same cues, such as promotions or word of mouth (Basuroy et al. 2006; Joshi and Mao 2012). Accordingly, we do not expect the popularity and artistic recognition of stars to differ in terms of how strongly they are associated with theatrical performance.

Assumption A common assumption among industry experts and academic researchers (issue 2 in Figure 1) predicts *fading star power*, such that stars' capacity to contribute to BO success has been on a downward trend in the motion picture industry overall (e.g., Merry 2015; *The Economist* 2016). According to this assumption, the diagnosticity of star brand equity must decrease with time. Regardless of fluctuations in their individual brand equity levels, if the association of stars with movie quality becomes less reliable, it follows that stars are less able to reduce moviegoers' choice uncertainty (Lynch et al. 1988). To clarify this question, we examine the evolution of movie quality variance over time. In an investigation of more than 10,000 movies released in the United States between 1950 and 2009 that received at least 1000 evaluations on IMDb, Johnston (2009) finds a positive and significant correlation ($r = .28$) between the year of release and users' ratings variance. Therefore, consistent with the fading star power assumption, we anticipate that the relationship of star brand equity with BO performance has been weakening over time.

Product reviews

Debate The relationship between professional critics' reviews and movie performance has been fiercely debated. According to an influencer argument, critics' reviews drive moviegoers'

choices (d'Astous and Touil 1999). A predictor view instead maintains that reviews reflect moviegoers' preferences but do not shape them (Basuroy et al. 2003; Zhang and Dellarocas 2006). This debate (issue 3 in Figure 1) also arises in studies that test the effects of critics' reviews on short- and long-term BO (Eliashberg and Shugan 1997). For example, because critics attend advance screenings, they often are the only objective sources of information in the early stage of a movie's life cycle (Litman 1983), and most professional reviews appear in the first week of a film's theatrical run (Legoux et al. 2016). Therefore, a movie's short-term BO revenue likely represents the best measure of the extent to which critics sway audiences, as influencers. However, after the initial release week, word of mouth about movie quality spreads, competing with the critical reviews to influence moviegoers. If critics are influencers, they should have significant impacts on short- but not on long-term BO. If they mirror audiences' taste, as predictors, this association should center on long- rather than short-term BO (Eliashberg and Shugan 1997). Alternatively, perhaps critics' reviews exhibit positive correlations with both short- and long-term BO. In this case, Basuroy et al. (2003) suggest that a decrease in the strength of the association between reviews and BO over a movie's theatrical run is consistent with an influencer role.

As this discussion implies, investigations of the short- and long-term effects of critics' reviews have produced inconsistent results, with varied support for the influencer role (Reinstein and Snyder 2005; Zuckerman and Kim 2003), the predictor role (Eliashberg and Shugan 1997), and a dual influencer–predictor role (Basuroy et al. 2003; Boatwright et al. 2007). From an accessibility–diagnosticity perspective, critics should primarily influence, rather than predict, movie success, because these experts possess credentialed authority, acquired from their formal training and institutional positions (Flanagin and Metzger 2013). Their reviews therefore should be diagnostic of movie quality, especially when competing cues are sparse (Lynch et al. 1988), such as immediately after the movie's release. In this sense, critics' reviews, as influencing factors, should relate significantly to short-term BO (Eliashberg and Shugan 1997). In later stages though, critics' reviews may be less diagnostic, because the proportion of the audience that has seen the movie is greater, and fewer moviegoers remain to be influenced (Basuroy et al. 2003). In addition, critics' reviews become less accessible, because they must compete with other cues that are readily available in the marketplace (Joshi and Mao 2012; Purohit and Srivastava 2001), such as word of mouth from friends who have seen the movie and share their

opinions in person or online (Liu 2006). Therefore, we expect the impact of critics' reviews on BO to be weaker over the long run, consistent with an influencer effect (Basuroy et al. 2003).

Assumptions According to the *users' advantage* assumption (issue 4 in Figure 1), consumers' reviews have a stronger impact on BO than critics' reviews do (Gazley et al. 2011; Holbrook and Addis 2007). Another common assumption predicts an *increasing users' advantage* (issue 5 in Figure 1), such that the impact of users' reviews on BO strengthens over time, while that of critics' reviews weakens (Bughin et al. 2010; Chakravarty et al. 2010; Duan et al. 2008; Flanagin and Metzger 2013). These assumptions largely stem from the sense that the arrival of ratings websites and social media spread users' reviews more widely (Proserpio and Zervas 2017), making them even more accessible than critics' reviews (Chen et al. 2012). In line with these assumptions, industry reports indicate that consumers trust other consumers' opinions more than other sources of information, including critical reviews (e.g., Nielsen 2015).

However, both empirical evidence and conceptual insights based on the accessibility–diagnosticity framework challenge such assumptions. First, in their meta-analysis, Floyd et al. (2014) show that experts have a stronger influence on product sale elasticities than do non-experts. Their study encompasses multiple product categories, but there is no theoretical reason to presume that their findings would not hold for movies. Second, these assumptions imply that users' reviews are somehow more diagnostic of movie quality than critics' reviews or else that users' reviews undermine the accessibility of critics' reviews (Lynch et al. 1988). Yet such effects are implausible, in that users' and critics' reviews differ by nature (Flanagin and Metzger 2013), as we noted previously. Both of them can reduce movie quality uncertainty, independently of each other. Therefore, we anticipate that the users' advantage and increasing users' advantage assumptions do not hold, considering their lack of empirical or conceptual foundation.

Methods

Study retrieval

Our methodological approach follows Watson et al.'s (2015) recommendations for meta-analyses. We used multiple sources to identify original studies that report empirical results,

available on or before September 1, 2015. First, we searched electronic databases—ScienceDirect, Scopus, Emerald, ProQuest (ABI/INFORM Global and dissertation abstracts), Academic Search Complete, Business Search Complete, JSTOR, and VERA—using the keywords "cinema," "film," "motion picture," "movie," "success," "performance," and "box office." Second, we manually reviewed the articles identified through this electronic search to uncover any additional references. Third, we conducted a manual shelf search of leading journals that often publish studies of motion picture performance across six academic disciplines that feature the most contributions to this topic: marketing, which included *Journal of Marketing, Journal of the Academy of Marketing Science*, *Journal of Marketing Research*, *Marketing Science*, *International Journal of Research in Marketing*, *Journal of Retailing*, and *Marketing Letters*; strategy and organizational theory, including *Academy of Management Journal, Academy of Management Review*, *Administrative Science Quarterly*, *Industrial and Corporate Change*, *Journal of Management*, *Organization Science*, and *Strategic Management Journal*; cultural economics, with *International Journal of Arts Management*, *Journal of Cultural Economics*, and *Journal of Media Economics*; sociology, featuring *American Journal of Sociology*, *American Sociological Review*, and *Technological Forecasting & Social Change*; and psychology, which included *Psychology of Aesthetics*, *Creativity and the Arts*, and *Psychology & Marketing*. Fourth, we physically visited three business school libraries and three prominent film libraries to search for publications: the Rosenfeld Library at UCLA (Westwood, CA), the London Business School Library (London, UK), the Myriam and J.-Robert Ouimet Library at HEC Montréal (Montréal, Canada), the Louis B. Mayer Library at the American Film Institute (Los Angeles, CA), the British Film Institute National Library (London, UK), and the Centre National du Cinema Library (Paris, France). Fifth, to address potential file drawer issues (Ferguson and Brannick 2012), we inquired after any unpublished manuscripts or forthcoming papers on motion picture performance on the Electronic Marketing List Information (ELMAR, ≈ 5,000 subscribers), Academy of Management (Business Policy and Strategy and International Management Division Lists, ≈ 3,300 and 1,600 subscribers, respectively), and the Association for Cultural Economics International (ACEI, ≈ 150 members) listservs. Sixth, we cross-checked our pool of primary studies with previous meta-analyses of star brand equity (Hofmann et al. 2017) and product reviews (Babić Rosario et al. 2016; Floyd et al. 2014; You et al. 2015).

A study was eligible for inclusion if it met two main criteria: First, movies, rather than consumers or studios, had to form the unit of analysis. Thus, studies by Sood and Drèze (2006), who ask respondents to evaluate movies, or Joshi and Hanssens (2009), who use the stock value of studio-owning companies as the dependent variable, were not included. Second, the study must include effect sizes (r) that indicate the relationship strength for the links of short-term BO, long-term BO, or other indicators of BO performance (e.g., theatrical admissions, international BO) with market, media, or artistic star brand equity or else with critics' or users' (volume and valence) reviews. We extracted some effect sizes directly, using the zero-order correlation reported, and others indirectly, by converting reported t- and p-values into Pearson's correlation coefficients using formulas provided by Borenstein (1994), Borenstein et al. (2009), and Lipsey and Wilson (2001) (see Online Appendix 1 for details). Several studies relied on the exact same sample of movies, so in these cases, we coded the study published earliest for effect sizes first, then added any new effect sizes from subsequent studies.

The search process yielded 643 effect sizes, but we removed 9 cases among 23 outliers that we identified, using several tests. That is, for each meta-analysis, we inspected any individual effect sizes that yielded either a standardized residual greater than 2.57 or a parameter change index above 1. We then excluded the outliers that could not be attributed to a small sample size (which increases sampling error and may yield false outlying values). The final sample therefore consists of 634 effect sizes from 150 studies over a four-decade span, including 21 (14%) that were not published (see Online Appendix 2)—a ratio superior to most marketing meta-analyses (Eisend and Tarrahi 2014). On average, there were 4.2 effect sizes per study (minimum = 1; maximum = 54), and marketing (31%), economics (26%), and management (12%) publications contributed the most effect sizes.

Study-level characteristics

We developed a coding scheme (Table 2) to identify moderators and control for methodological differences across studies (the Appendix contains the correlations among these variables). Studies of the popularity or artistic recognition of stars examine one of three star brand equity dimensions: market, based on the BO of the stars' previous movies; media, based on whether stars appear in popular media lists (e.g., *Empire*, *Variety*); and artistic, based on nominations and awards bestowed by the industry (e.g., Academy Awards, Screen Actors' Guild, Golden Globes).

In addition, stars may be actors, directors, other contributors (e.g., music composers, screenplay writers), or a mix of these different occupations. Product reviews originate from professional critics (e.g., rottentomatoes.com, *Variety*) or amateur users (e.g., Yahoo!Movies, Twitter). To test the fading star power and increasing users' advantage assumptions, we coded each effect size for the median year of release of the movies sampled in the primary study and the range of release years.

[Insert Table 2 about here]

We also coded several control variables according to whether the primary studies used convenience or random sampling, whether they controlled for movies that won or were nominated for awards, and the geographic market. In addition, we determined whether the effect of any given movie quality cue depended on the presence of other cues, in two ways. First, we dummy coded for whether each effect size was a zero-order or partial correlation, to control for the presence of other drivers of BO performance in the primary study model (e.g., promotion, genre, production budget, sequel). Second, we added specific dummy variables to indicate whether each effect size came from a model that also included other focal effect sizes, namely, the artistic, market, and media dimensions of star brand equity and critics' and users' reviews. With these two controls, we can better model the actual accessibility of the focal cues. We also coded for whether the effect size came from a top-tier journal or not, to account for study quality. Finally, to control for potential publication bias, we noted the precision of the effect size, as the inverse of its standard error (Stanley and Doucouliagos 2012), as well as whether the study was published or unpublished.

Between-study movie overlap

The primary studies' units of analysis (i.e., movies) represent a well-defined population. Since 1930, approximately 53,000 U.S.-produced full feature films have been released in theaters (Internet Movie Database 2016). Considering the number of primary studies that rely on this population, the potential for movie overlap is not negligible. For example, in 1998, 1,225 U.S.-produced films were released in theaters; four primary studies (sample sizes ranging from 108 to 1,542) examined the relationship of the artistic dimension of star brand equity and long-term BO using sampling frames that included that year. The study with 1,542 movies spanned 15 years; hence, we assumed that it included 102.8 of the 1,225 movies from 1998 (1,225/15). Applying

the same logic to all these studies, we determine that 295 movies were sampled by studies that reported an artistic–long-term BO effect size, from the 1,225 movies released in 1998. In this scenario, it may be unlikely that the mean effect size reflects the exact same movies, but some movies likely overlap across the four primary studies, which could increase the potential for Type I error (Lin and Sullivan 2009).

To circumvent this issue, we first applied the method suggested by Munder et al. (2013) at the summary effect size level (e.g., association between artistic star brand equity and long-term BO). This method spreads overlapping movies across the primary studies and adjusts their sample size, according to an assumption that each movie appears only once across all studies for a given summary effect size. We also reproduced all the summary effect sizes in Table 3 with a subsampling approach (Mukhopadhyay 2008) that features only individual effect sizes from studies with non-overlapping years. This step provides an estimate of the extent to which movie overlap might affect the results. The first procedure informs all our subsequent analyses, such that we use the adjusted sample size for all individual effects. The second procedure confirms that the subsampled summary effect sizes are similar to those from the full sample (for details, see Online Appendix 3). No previous meta-analyses that use primary studies to investigate movies have accounted for this potential overlap bias (Babić Rosario et al. 2016; Floyd et al. 2014; Hofmann et al. 2017; You et al. 2015).

Analytical approach

The analyses all were performed with the 1.9-8 "Metafor" package in the R software, 3.1.1 version, using inverse variance-weighted effect sizes and restricted maximum likelihood estimators (Viechtbauer 2010). In a series of univariate meta-analyses for all the relevant bivariate relationships, the independent variables were market star brand equity, media star brand equity, artistic star brand equity, critics' reviews valence and volume, and users' review valence and volume; the dependent variables were short-term BO, long-term BO, and other performance measures. Table 3 displays the 21 mean effect sizes, along with their confidence and prediction intervals and heterogeneity and publication bias indicators. To test the focal debates and assumptions, we performed a series of meta-regressions that included all the moderators, to account for their simultaneous effects and possible correlations (Borenstein et al. 2009; Hunter and Schmidt 2004). These analyses relied on Fisher's z-scores, transformed from the correlations, then back-transformed to r to indicate the mean effect sizes and confidence and

prediction intervals (Borenstein et al. 2009; also see Online Appendix 1). The same analyses on untransformed correlations yielded similar results (available on request), with the same substantive interpretations.

Model development

Independence in the effect sizes is not warranted; the 634 effect sizes are embedded within 150 articles. Thus, we must account for the nested structure of the data (Bijmolt and Pieters 2001). In addition, we fit a random effects, rather than fixed effect, model, as recommended by Ferguson and Brannick (2012), because our objective was to generalize the results to the population of movies,[3] rather than find mean effect sizes common to the primary studies. We thus applied a multilevel meta-analytic approach (Van den Noortgate et al. 2015) with two random effects, at the effect size and article levels, to decompose total heterogeneity. Formally, the model we use to estimate the Table 3 summary effect sizes is:

$$ES_{ij} = \beta_0 + \upsilon_{0j} + \psi_{ij +} + e_{ij}, \tag{1}$$

where ES_{ij} reflects the measure of effect size, β_0 is the overall effect size without moderators, υ_{0j} indicates the estimate of the extent to which the effect sizes within articles correlate, ψ_{ij} represents between-effect size variance, and e_{ij} is the within-effect size variance (sampling error). The full star brand equity model thus is (see the variable operationalizations in Table 2):

$ES_{ij} = \beta_0 + \beta_1$ random sampling $+ \beta_2$ other market $+ \beta_3$ US market $+ \beta_4$ artistic dimension $+ \beta_5$ media dimension $+ \beta_6$ other performance measures $+ \beta_7$ short-term BO $+ \beta_8$ median year of movie release $+ \beta_9$ range of movie release years $+ \beta_{10}$ actor star power $+ \beta_{11}$ director star power $+ \beta_{12}$ other contributors star power $+ \beta_{13}$ movie awards $+ \beta_{14}$ published $+ \beta_{15}$ top-tier publication $+ \beta_{16}$ zero-order correlation $+ \beta_{17}$ effect size precision $+ \beta_{18}$ artistic dimension included $+ \beta_{19}$ market dimension included $+ \beta_{20}$ media dimension included $+ \beta_{21}$ critics' reviews included $+ \beta_{22}$ users' reviews included $+ \beta_{23}$ artistic dimension $\times$ other performance measures $+ \beta_{24}$ media dimension $\times$ other performance measures $+ \beta_{25}$ artistic dimension $\times$ short-term BO $+ \beta_{26}$ media dimension $\times$ short-term BO $+ \beta_{27}$ artistic dimension $\times$ median year $+ \beta_{28}$ media dimension $\times$ median year $+ \upsilon_{0j} + \psi_{ij} + e_{ij}$.

$$(2)$$

The product reviews model is:

$ES_{ij} = \beta_0 + \beta_1$ random sampling $+ \beta_2$ other market $+ \beta_3$ US market $+ \beta_4$ users' review $+ \beta_5$ review volume $+ \beta_6$ other performance measures $+ \beta_7$ short-term BO $+ \beta_8$ median year of movie release $+ \beta_9$ range of movie release years $+ \beta_{10}$ movie awards $+ \beta_{11}$ published $+ \beta_{12}$ top-tier publication $+ \beta_{13}$ zero-order correlation $+ \beta_{14}$ effect size precision $+ \beta_{15}$

---

[3] According to the between-study movie overlap analysis, few years were oversampled by the primary studies, so not all movies released were covered. Generalizing our results remains an important objective, despite the relatively smaller population compared with other meta-analysis projects.

artistic dimension included + β_{16} market dimension included + β_{17} media dimension included + β_{18} critics' reviews included + β_{19} users' reviews included + β_{20} users' review × review volume + β_{21} users' review × median year of movie release + $\upsilon_{0j} + \psi_{ij} + e_{ij}$.

(3)

Results

Effect size distribution heterogeneity

For each summary effect size in Table 3, we compute three indicators of heterogeneity: tau^2, which is the sample estimate of the between-effect size variance and thus the amount of effect size heterogeneity that can be accounted for by moderators; I^2, reflecting the between-effect sizes variance, as a percentage of total observed variance; and the Q-statistic that indicates the effect size dispersion. The random effects model is appropriate for all summary effect sizes; they exhibit a significant Q-statistic and I^2 values greater than 50% in all cases except for users' valence–short-term BO[4] (Borenstein et al. 2009; Hunter and Schmidt 2004). The results in Table 4 also reveal that the meta-regressions explain substantial between-effect sizes variance (> 90%), in support of the inclusion of moderators (Borenstein et al. 2009). For all the meta-regressions, we verified that the variance components tau^2 and pho (i.e., sample estimate of the correlation between effect sizes within articles) could be identified; we inspected whether the parameters corresponded with the peaks of their log-likelihood profile (Konstantopoulos 2011).

[Insert Table 3 and Table 4 about here]

Publication bias

To address the threat of publication bias when estimating the summary meta-analyses (Table 3), we leveraged three indicators. First, we performed Egger's (Sterne and Egger 2005) regressions for each meta-analysis, for which a significant z-value suggests publication bias. This test is based on the ratio of the effect size by its standard error, regressed on the standard error, so a significant result indicates asymmetry in the funnel plot (Sterne and Egger 2005). Second, we computed both Rosenthal's (1979) fail-safe N, which indicates how many unavailable, null effect sizes would be required to bring the mean effect size below a .05 significance level (Borenstein et al. 2009), and Rosenberg's (2005) fail-safe N, which estimates the number of missing studies required to yield an insignificant weighted, rather than unweighted, mean effect size. When N is

---

[4] The sampling error for the users' valence–short-term BO effect size was greater than 50%, but the random effects model still was adequate, because the null hypothesis of distribution homogeneity was rejected (Q-statistic [13]) = 44.2, $p < .001$.

14

smaller than k, the effect size estimate is not robust (Ferguson and Brannick 2012). Third, we implemented Duval and Tweedie's (2000) trim-and-fill approach, which identifies effect sizes that create asymmetry in a funnel plot of their precision by their strength. They are then deleted (trimmed), allowing for the estimation of a new average effect size and its distribution, to which new effect sizes are added (filled) to improve the funnel plot's symmetry. Finally, this approach returns updated mean effect sizes, distributions, and heterogeneity estimates.

These indicators (Egger's regression, fail-safe Ns, trim-and-fill) are prone to Type I error, so we also used a tandem approach (Ferguson and Brannick 2012) and deemed publication bias likely if it was suggested by at least two indicators. If a summary effect size thus appeared potentially affected by publication bias, we checked if the tests of the debates and assumptions to which it related were robust to the trim-and-fill–adjusted mean effect size. However, this adjustment cannot be designated the "true" effect size, because asymmetry is not necessarily due to publication bias (Ferguson and Brannick 2012). Therefore, we regard a non-adjustment of the effect size as a more reliable indicator of the absence of publication bias, relative to an adjustment as an indicator of the publication bias–corrected effect size.

Finally, the meta-regression models account for publication bias in two ways. We include moderators that indicate whether a particular effect size comes from a published or unpublished manuscript. We also control for effect size precision, with the inverse of the standard errors (Stanley and Doucoulagios 2012). Small sample studies often are published only if their effect sizes are large enough (Ferguson and Brannick 2012; Kepes et al. 2012).

Star brand equity

Star popularity versus artistic recognition debate We investigate issue 1 in Figure 1 by comparing the market and media dimensions of star brand equity (star popularity) with its artistic dimension (artistic recognition). Our results partly confirm our prediction that the popularity and artistic recognition of stars do not differ in the strength with which they drive theatrical performance. As we show in Table 5, Model 1, the artistic dimension of star brand equity does not differ from the market dimension, according to the non-significant main effect (β = -.040, SE = .038, ns), a result that contrasts with Hofmann et al.'s (2017) finding that the market dimension of star brand equity yields larger effect sizes than its artistic dimension. In a supplementary

analysis, in which we use the media dimension, rather than the market one, as a reference category, we confirm that the media dimension is not different from the artistic dimension ($\beta = -.033$, SE = .038, *ns*). These results are available on request.

We also estimate the same model after adding the interaction parameters between star brand equity and the movie theatrical performance measures, as well as between star brand equity and the median year of movie release. Using the media dimension as the reference category, none of the interactions of brand equity and performance produce significant parameters (z-values < 1.35, ns; results available on request). However, when the reference category is the market dimension (Table 5, Model 2), the interaction between the artistic dimension and short-term BO is significant, though only at $\alpha = .10$, rather than the conventional $\alpha = .05$ level ($\beta = -.089$, SE = .046, $p < .10$). Still, we probe this interaction to establish a comparison with Hofmann et al. (2017). The intercept-free meta-regression that controls for the continuous moderators indicates that the market dimension has a greater effect than the artistic dimension on short-term BO (artistic $\bar{r} = .034$, k = 6; market $\bar{r} = .155$, k = 41; $\chi^2(1) = 8.30$, $p < .01$); their impacts on long-term BO are similar (artistic $\bar{r} = .104$, k = 24; market $\bar{r} = .151$, k = 62; $\chi^2(1) = 3.11$, *ns*). The same model, applied to a subset of effect sizes that came from studies that estimated both artistic and market dimensions (in order to hold study characteristics and model specifications constant; see Online Appendix 4, column 1), reveals that the main effect of the artistic dimension is not significant ($\beta = .008$, SE = .061, *ns*), nor is its interaction with short-term BO ($\beta = .056$, SE = .065, *ns*). Another analysis, for the subset of effect sizes that came from studies that examined both artistic and media dimensions, also yields a non-significant main effect for the artistic dimension ($\beta = -.018$, SE = .028, *ns*).[5] In summary, these increasingly specific analyses contest the notion that star popularity is superior to artistic recognition in its effects (cf. Hofmann et al. 2017).

The presence of star actors yields the same effect as the presence of star directors (Wald-type z-value = -.71, *ns*). This finding alleviates concerns that our results differ from Hofmann et al.'s (2017) simply due to the inclusion of star directors and other star contributors. Nevertheless, Hofmann et al.'s results could hold if the artistic recognition of star directors had a stronger

---

[5] The total number of effect sizes that we could include in this latter model is limited (k = 13; artistic = 6 and media = 7), and only the effects of the continuous parameters and the artistic dimension could be estimated. The full model results are available on request.

effect than that of star actors for long-term BO. To test for this possibility and directly compare star actors and directors, we ran a meta-regression similar to Model 1 in Table 5 that includes effect sizes related solely to actors or directors as contributors (see Online Appendix 4, column 2). The three-way interaction among director (with actor as the reference category), short-term BO, and star brand equity dimensions is not significant (z-values < 1.22, *ns*), so it is unlikely that our conclusions diverge from Hofmann et al.'s (2017) simply due to the inclusion of star directors. Our model also exerts a tighter statistical control than in Hofmann et al. (2017), in that our meta-regressions account for publication bias (i.e., manuscript status, effect size precision), the inclusion of other focal effect sizes in the primary studies' model, and whether the primary studies controlled for movie quality. All our analyses rely on adjusted sample sizes too, to account for movie overlap in the primary studies.

[Insert Table 5 about here]

Sensitivity analysis These findings are robust to effect size distribution asymmetry. That is, we performed Wald-type tests on the trim-and-fill–adjusted mean effect sizes and their standard errors (last column of Table 3) if the tandem procedure indicated possible publication bias (Ferguson and Brannick 2012). The adjusted market dimension and short-term BO effect size continues to emerge as greater than that of the unadjusted artistic and short-term BO effect (z-value = -2.52, *p* < .05). The adjusted media dimension and short-term BO effect size is not significantly different from the unadjusted artistic and short-term BO effect size (z-value = .31, ns), confirming our prediction of this debate. In summary, we find no pertinent difference between the effects of star popularity and artistic recognition on long-term BO. However, the market dimension of star popularity (but not the media dimension) has a stronger effect on short-term BO than does the artistic dimension of star recognition.

Fading star power assumption To address the major assumption in this category, we ask: Is the impact of stars on movie performance weakening over time (issue 2, Figure 1)? As we anticipated, the answer is yes, but only with regard to star popularity. The significant, negative, main effect of the median year of movie release in Model 1 illustrates that the impact of the market dimension of star brand equity on long-term BO has weakened over time (β = -.002, SE = .001, *p* < .05). But in Model 2, the significant, positive interaction of the artistic dimension of star brand equity with the median year suggests it still is more stable than the market dimension

17

over time (β = .005, SE = .002, p < .01). Using the media dimension as the reference category instead of the market dimension yields similar results: median year has a negative main effect (β = -.004, SE = .002, p < .05), and the media dimension is losing its power over time relative to the artistic dimension (interaction of median year with the artistic dimension: β = .006, SE = .002, p < .01; results available on request). Overall, this finding supports the fading star power assumption for star popularity but not for artistic recognition, in stark contrast with Hofmann et al.'s (2017) finding that the impact of star power does not erode with time. Even though both studies operationalize the impact of time as the median year of movie release and control for the years spanned by the sampled movies, this discrepancy is plausible: We adopt tighter statistical controls and consider more effect sizes (303 versus 172). Furthermore, we allowed for an estimate of the interaction parameter between brand equity dimensions and time, which was absent from Hofmann et al.'s (2017) study.

A closer examination reveals that the effect sizes of the market and media dimensions once were significantly greater than those of the artistic dimension, but that is no longer the case. In Figure 2, Panels a and b, we plot the regression lines of the relationship between the median year of movie release in the primary studies and the effect sizes for the market and media dimensions, respectively.[6] A reference (red) line representing the mean effect size for the artistic dimension provides a constant. The market and artistic dimensions do not differ significantly since 1940 (i.e., when the confidence bands around the regression line cross the reference line), and their mean effect sizes have been of the same magnitude since 2000 (i.e., when the regression line crosses the reference line). The decline of the market dimension has been steep though. Its mean effect size decreased from $\bar{r}$ = .245 in the pre-World War II period to about $\bar{r}$ = .050 in the 2010s; after 1980, it no longer was significantly different from 0. Similar findings apply to the media dimension. Its confidence band crosses the artistic effect size since the mid-1920s, with similar mean effect sizes since the early 2010s.

[Insert Figures 2a and 2b about here]

To establish this negative effect of time for star popularity, we provide forest plots of the individual effect sizes for market brand equity, sorted from oldest to most recent, in Figure 3,

---

[6] To do so, we use meta-regressions to estimate which moderators from Model 2 are significant for the market and media dimensions. Then, with Johnson and Huedo-Medina's (2011) procedure, we "move" the model's intercept across the values of the significant moderator(s) to compute a series of confidence intervals. Dummy variables for categorical moderators are orthogonally coded (-1; 1).

Panels a (1935–1999; k = 61) and b (2000–2013, k = 70). Each effect size is represented by a square, proportional to its weight in the mean effect size, and the lines around the square representing the 95% confidence interval (CI). In Panel a, the most precise effect sizes (i.e., narrowest 95% CI) are clear of the significance line, but in Panel b, they are not significant. The mean effect size in Panel a ($\bar{r}$ = .168, 95%CI = .10–.23) is significantly greater than that for Panel b ($\bar{r}$ = .080, 95% CI = .04–.11; z-value = 2.32, $p < .05$). Rather than cue diagnosticity, an alternative explanation for this result might be that the strength of star popularity weakened in the legal environment that became less favorable to studios after 1948.[7] However, a comparison of average effect sizes before and after 1948 is not significant (z-value = 1.00, *ns*),[8] offering confidence in our explanation that the loss of star popularity diagnosticity is due to increasing variance in movie quality.

[Insert Figures 3a and 3b about here]

Hofmann et al. (2017) find that movies released in the United States exhibit greater effect sizes than movies released elsewhere whereas the market variables in Model 2 are not significant. We had enough effect sizes to make meaningful comparisons across three categories, namely, the United States, Europe, and other regions, rather than just U.S. versus non-U.S. releases. Accordingly, we confirm that the non-significant effect is not dependent on the choice of reference category. When we re-estimate Model 2 by specifying "other" as the reference category and then by collapsing other and Europe into one category, the results remain unchanged in both cases (Online Appendix 4, columns 3 and 4).

Product reviews

Influencer versus predictor debate To address the debate surrounding the role of professional critics as influencers or predictors (issue 3, Figure 1), we first ran an intercept-free meta-regression on the effect sizes of critics' reviews, to obtain mean estimates for the combinations of the review metric and performance indicator moderators. The effect of reviews on short- and long-term BO is positive and significant for both volume ($\bar{r}$ = .272, k = 10, $p < .001$; $\bar{r}$ = .344, k = 16, $p < .001$, respectively) and valence ($\bar{r}$ = .148, k = 71, $p < .001$; $\bar{r}$ = .216, k = 67, $p < .001$,

---

[7] In 1948, federal antitrust law prevented studios from engaging in two previously common practices: signing low-wage creative personnel to long-term contracts and vertically integrating film distribution (Balio 1985).
[8] We excluded 2000–2013 to compare the pre- and post-1948 eras without contamination by competing cues, such as users' reviews, which became far more widespread after 2000. Including this period yields similar results (available on request).

respectively). These results are consistent with the notion that critics can be both influencers and predictors (Eliashberg and Shugan 1997). However, if the impact of critics' reviews is weaker for long-term versus short-term BO (Basuroy et al. 2003), our data still might prioritize an influencer role. Another meta-regression (Model 3, Table 6) shows that critics' reviews have a marginally weaker effect for short-term BO than for long-term BO (β = -.053, SE = .027, p < .10), ruling out this influencer-only role (Basuroy et al. 2003). In summary, and partly in line with our prediction, these analyses are compatible with a dual influencer–predictor role for critics.

Sensitivity analysis For a more rigorous test of the role of critics' reviews on BO performance, we ran the same model for a subsample of effect sizes obtained only from studies that examine both short- and long-term BO. The intercept-free meta-regression reveals positive, significant, short- and long-term BO effect sizes for volume ($\bar{r}$ = .278, k = 10, p < .001; $\bar{r}$ = .396, k = 9, p < .001, respectively) and valence ($\bar{r}$ = .144, k = 27, p < .001; $\bar{r}$ = .241, k = 26, p < .001, respectively). The full meta-regression leads to the same conclusion as the full sample results (Online Appendix 5, column 4). The impact of reviews is weaker for short- than for long-term BO (β = -.089, SE = .039, p < .05), regardless of the metric (interaction with volume β = -.030, SE = .081, ns). This result confirms that critics act as both influencers and predictors.

Users' advantage assumption As anticipated, we find little support for the users' advantage assumption in relation to the volume of reviews (issue 4, Figure 1). In Model 4 of Table 6, we find no main effect of users' versus critics' reviews (β = -.027, SE = .034, ns); however, the interaction with the volume of reviews is significant at α = .10 (β = .094, SE = .048, p < .10). If we control for the continuous moderators, the intercept-free meta-regression indicates that the *valence* of product reviews yields similar effect sizes for users ($\bar{r}$ = .206, k = 82) and critics ($\bar{r}$ = .212, k = 172; $\chi^2(1)$ = .05, ns) but the *volume* of users' reviews has a stronger effect size ($\bar{r}$ = .392, k = 45) than that of critics' reviews ($\bar{r}$ = .297, k = 32; $\chi^2(1)$ = 4.69, p < .05). Whereas the volume of reviews has a positive main effect (β = .084, SE = .034, p < .05), its dominance over valence appears greater for users ($\bar{r}$ = .392 versus $\bar{r}$ = .206; $\chi^2(1)$ = 27.83, p < .05) than for critics ($\bar{r}$ = .297 versus $\bar{r}$ = .212; $\chi^2(1)$ = 5.86, p < .05).

Sensitivity analysis Primary studies including the effect sizes of both users' and critics' reviews could provide a more thorough test of the user's advantage assumption.[9] Therefore, we applied Model 4 to a subsample of studies that investigated both users and critics (k = 149). As shown in Online Appendix 5 (column 1), the main effect of users' reviews (β = .0005, SE = .030, *ns*) is not significant, which mirrors the full sample results. However, the interaction with review metric (volume vs. valence) is not significant here (β = .056, SE = .059, *ns*). Critics' and users' reviews have similar effect sizes, whether we consider valence ($\bar{r}$ = .169, k = 53 versus $\bar{r}$ = .175, k = 53, respectively; $\chi^2(1)$ = .18, *ns*) or volume ($\bar{r}$ = .234, k = 28 versus $\bar{r}$ = .277, k = 18, respectively; $\chi^2(1)$ = .87, *ns*). Comparing critics' and users' reviews using effect sizes obtained in similar conditions thus refutes the users' advantage assumption.

In contrast, estimating Model 4 with subsamples of studies that investigate the effect sizes of either users' or critics' reviews only (column 2) yields results that are consistent with the full sample (main effect: β = -.033, SE = .052, *ns*; interaction: β = .164, SE = .082, *p* < .05; volume: $\bar{r}$ = .541, k = 14 versus $\bar{r}$ = .341, k = 17, respectively; $\chi^2(1)$ = 6.35, *p* < .01; valence: $\bar{r}$ = .224, k = 119 versus $\bar{r}$ = .229, k = 29, respectively; $\chi^2(1)$ = .01, *ns*). Ultimately, the similar effects for valence across sampling conditions and the lack of robust superiority for volume suggest that support for the users' advantage assumption is weak. The greater difference between the effect sizes of volume and valence when it comes to users', rather than critics', reviews remains robust though, observed in the full sample and confirmed by the analysis of those studies that include both types of product reviews. In this latter subsample, the volume and valence of critics' reviews exhibit similar strength ($\bar{r}$ = .234 versus $\bar{r}$ = .169, respectively; $\chi^2(1)$ = 1.51, *ns*); the volume of users' reviews outperforms their valence ($\bar{r}$ = .277 versus $\bar{r}$ = .175; $\chi^2(1)$ = 2.53, *p* < .05).

Increasing users' advantage assumption This last assumption (issue 5, Figure 1) stipulates that the impact of users' reviews on BO performance has increased, while that of critics' has decreased. As anticipated, our results do not support this assumption. In Model 4 in Table 6, neither the main effect of the median year of movie release (β = -.003, SE = .003, *ns*) nor its interaction with users' reviews (β = -.003, SE = .006, *ns*) is significant.

---

[9] We thank an anonymous reviewer for suggesting this analysis.

Sensitivity analysis A direct effect of critics' reviews would manifest as an influencer effect (Eliashberg and Shugan 1997), so we ran a meta-regression (Online Appendix 5, column 3) for a subsample of short-term BO effect sizes only, as a stricter test of the increasing users' advantage assumption. The main effect of users' reviews ($\beta = -.041$, SE = .061, ns), the main effect of the median year of movie release ($\beta = .007$, SE = .007, ns) and its interaction with users' reviews ($\beta = -.023$, SE = .087, ns) are not significant, so we confirm the Model 4 results.

[Insert Table 6 about here]

Figures 4 and 5 summarize the findings detailed throughout this Results section.

[Insert Figures 4 and 5 about here]

Discussion

Moviegoers use accessible, diagnostic cues to avoid the risk of purchasing a ticket to a cinematographic equivalent of a "lemon" (Akerlof 1970). Star brand equity and product reviews offer credible sources of such cues and thus have been the focus of hundreds of academic studies over the past 40 years, producing 634 individual effect sizes that we can analyze. In doing so, we provide an empirical test of the cue accessibility–diagnosticity framework (Lynch et al. 1988), which confirms most of the predictions that we derived from this framework. The findings are particularly relevant to industries that market experiential products whose quality cannot be ascertained prior to consumption (Eliashberg and Sawhney 1994). By synthesizing primary studies of star brand equity and product reviews, our investigations also make several contributions to cinema-related literature.

First, contrary to Hofmann et al.'s (2017) assertion that the market dimension of brand equity relates more strongly to BO performance than its artistic dimension, we specify that this finding holds for the initial release period only (e.g., opening weekend). When we consider BO revenue over the movies' entire theatrical lifespan, the market and artistic dimensions exhibit similar impacts. This result adds crucial nuance to the debate about the superiority of the popularity or artistic recognition of stars in terms of driving BO success.

Second, our more fine-grained approach to the operationalization of star popularity reveals that the media dimension of star brand equity does not have a stronger impact on BO than its

artistic dimension. The effect of star popularity on BO performance, identified by the market and media dimensions, actually is in sharp decline, to the extent that its current impact is null. In contrast, the effect of artistic recognition on BO performance has remained small ($\bar{r} = .072$) but steady over time. Distinguishing star popularity from artistic recognition leads to different results than Hofmann et al.'s (2017) finding that the impact of star brand equity has remained constant over time. In this sense, our research provides new support for the fading star power assumption. Furthermore, in follow-up analyses (forest plots in Figure 3), we exclude alternative explanations, such as the changing legal environment post-1948 that might have weakened popular stars' impact. Rather, increased variability in the quality of movies in which stars feature seemingly has hampered their diagnosticity (Lynch et al. 1988). This reasoning receives empirical support from the increasing variance of users' ratings of more than 10,000 U.S. movies released since 1950 (Johnston 2009).

Third, we provide comprehensive empirical insights into the influencer–predictor debate related to the role of critics. Both the volume and the valence of critics' reviews exhibit significant positive relationships with short- and long-term BO, suggesting they function as both influencers and predictors (Eliashberg and Shugan 1997). Consistently, we do not find a greater impact of critics' reviews on short- versus long-term BO, which rules out an influencer-only effect (Basuroy et al. 2003). The effect size of critics' reviews instead is greater for long-term BO, so their predictor role even might be more prominent. This finding is only partially consistent with the accessibility–diagnosticity framework though, in that it suggests an influencer role. In the wake of Eliashberg and Shugan (1997) and Basuroy et al. (2003), we acknowledge that a comparison of short- and long-term BO cannot resolve the debate definitively; movie choices always involve individual and situational factors that BO revenue studies cannot capture. However, Eliashberg and Shugan (1997) and Basuroy et al. (2003) ignore the potential role of critics as predictors of movie success initially or influencers of consumption choices later. They simply highlight the role that appears most compatible with the pattern of results across short- and long-term BO. Our study adds to this discussion, by suggesting that the impact of critics is not limited to a direct effect on moviegoers' decisions in the first week of the film's lifecycle. Critics appear able to tease out the intrinsic qualities that reflect moviegoers' preferences, as well as the intensity of word of mouth about the movie throughout its theatrical run. Part of the effect

identified through our meta-analysis may thus stem from the influence of critics on exhibitors'
decisions to continue screening or pull a movie from their screens (Legoux et al. 2016).

Fourth, our study provides a first comprehensive test of the conventional wisdom that the
impact of users' reviews looms larger than that of critics' (e.g., Bughin et al. 2010; Chakravarty
et al. 2010; Chisholm et al. 2015). Our results refute this users' advantage assumption, in terms
of both valence and volume. Among existing meta-analytic investigations of product reviews,
only Floyd et al. (2014) address this assumption; they show that users have a credibility
disadvantage relative to expert reviewers. However, their results are restricted to sales elasticity
effect sizes, such that they consider only 4 primary studies about movies, versus 112 in our
study. Their study includes a range of product categories, so their finding may suggest that
experiential credibility is relatively more important than expertise for movies, relative to other
products (Flanagin and Metzger 2013). In addition, we find no evidence in support of the
increasing users' advantage assumption. The relationships of both critics' and users' reviews
with movie performance are stable, in terms of strength, over time, consistent with Chen et al.'s
(2012) claim that studio shareholders value the information provided by professional critics,
even in the Internet age. Users' reviews may be more widely available (Bughin et al. 2010; Liu
2006), but because users and critics have distinct roles in the motion picture ecosystem, due to
their different forms of credibility (Flanagin and Metzger 2013), their diagnosticity is
independent from each other's accessibility (Akdeniz et al. 2013; Joshi and Mao 2012; Lynch et
al. 1988; Purohit and Srivastava 2001). The assertion that "everybody's a critic" may be true
(Kermode 2014), but it has not altered the role of professional critics for the movie industry.

Our results also inform other research areas. For example, the recognition that review volume
has a stronger impact on performance than review valence for users but not for critics contributes
to broader literature on product reviews. Previous meta-analyses diverge in their conclusions;
based on elasticities, Floyd et al. (2014) and You et al. (2015) suggest valence has a superior
impact on sales, but using correlations, Babić Rosario et al. (2016) find that volume is the
superior metric. Correlations might be a more appropriate effect size than sales elasticities,
because they are more comparable across metrics (e.g., IMDb uses a 10-point scale, but
CinemaScore uses an A+ to F scale). Our results consolidate Babić Rosario et al.'s (2016)
reversal of Floyd et al.'s (2014) and You et al.'s (2015) conclusions, for the case of movies.
Furthermore, our findings indicate that the stronger effect of the volume of users' reviews,

compared with their valence (Babić Rosario et al. 2016), does not extend to critics' reviews. Bandwagon effects might explain some of the primacy of volume over valence for driving product sales (e.g., Van den Bulte and Lilien 2011), in that awareness of others' product judgments reduces uncertainty about quality, regardless of whether the judgments are positive or negative, because following others' behaviors (i.e., to consume a product) can be a less risky choice (Babić Rosario et al. 2016). Our results suggest that critics' reviews are not subjected to this bandwagon effect though, in further support of our contention that users' and critics' reviews offer cues with different natures (Flanagin and Metzger 2013).

Our findings also likely apply to other experiential products, such as books, music, video games, live shows, fine dining, travels, or education (e.g., Eliashberg and Sawhney 1994; Hennig-Thurau et al. 2015), for which ingredient brands, opinions of experts, and reviews by other consumers provide important cues. For example, the deleterious effect of product quality variability on the brand equity of popular stars likely applies to musicians, who produce albums that can polarize fans and choose whether to play former hits or new songs on tour. These results also can inform marketing initiatives that seek to associate stars with specific product or brands. Celebrity endorsements of low quality products may reduce the diagnosticity of the cues they provide consumers (Knoll and Matthes 2017), which could undermine their brand equity and ultimately make them less effective endorsers.

We acknowledge that our findings about the impact of expert reviews do not generalize to all hedonic products. Rather, we predict that they apply to products that feature quality uncertainty—that is, products with hedonic and experiential attributes. Cue diagnosticity theory concurs that it is more important to lower quality uncertainty for hedonic-experiential than for hedonic products (Lynch et al. 1988). Accordingly, Babić-Rosario et al.'s (2016) estimate of the mean correlation between product review valence and sales of hedonic products in general ($\bar{r} = .05$) is understandably smaller than the mean effect sizes of valence on short- ($\bar{r} = .102$, 95% CI $= .049, .155$) and long-term ($\bar{r} = .202$, 95% CI $= .163, .240$) BO that we find.

Finally, in terms of methodological contributions, our meta-analysis is unique, in that it accounts for overlap bias in a motion picture setting, as is customary in meta-analyses of units of analysis that represent relatively smaller populations. We adapt the method suggested by Munder et al. (2013) at the summary effect size level and reproduce all the summary effect sizes with a

subsampling approach (Mukhopadhyay 2008), so only individual effect sizes from studies with non-overlapping years serve to calculate the mean effect sizes. These additional precautions increase the robustness of our analysis; we hope they also spur other meta-analysts to account for potential overlap, such as when they investigate topics by using available company or industry secondary data.

Managerial implications

This meta-analytic investigation provides cinema executives with robust evidence that they can use to make more informed decisions. In particular, the fading power of star popularity, especially when considered in combination with evidence that movie stars' favorability ratings decay quickly (Luo et al. 2010), calls the bankability of stars into question. When they go to cast their most promising blockbusters, movie studios often rely on Q scores for actors, which rely on surveys of representative samples of a population to assess the familiarity and appeal of celebrities. Recent anecdotal evidence, such as Matt Damon's lead role in Zhang Yimou's 2016 movie *The Great Wall*, based on his popularity with Chinese audiences, aligns with our findings: three months after its release, this movie was projected to lose in excess of $75 million (McClintock and Galloway 2017).

If popularity and artistic recognition have equivalent impacts on total BO revenue, then casting A-listers is more effective to ensure movie success in the first week of the theatrical release. Risk-averse studio executives who seek to recoup their movie investments quickly thus may see popular stars as a sensible choice. The artistic recognition component of star brand equity is a stable predictor of BO over time, which also increases the importance of prizes such as Academy Awards. Despite diminished audiences for the broadcasts of award shows (Patten 2017), a star who has achieved artistic recognition appears to provide a diagnostic and accessible cue that is worth promoting. Popular stars should refrain from appearing in poor quality movies, because doing so lowers the diagnosticity of their brand equity. Relying on lesser known stars or B-listers, or investing in attributes other than human brands, such as special effects or appealing animated characters, thus may be a better strategy for movies with limited artistic aspirations.

Considering that critics take both influencer and predictor roles through the volume and valence of their reviews, they constitute both early indicators of how a movie will fare over its theatrical life and resources to improve the movie's performance. Determining what causes

critics to review a movie (or not) thus is as important as their average evaluations. For users, the volume of reviews is even more pertinent than their valence, such that it is an important barometer of film performance. In a related note, the lack of clear support for the users' advantage assumption suggests that marketers should be cautious before allocating more promotional effort toward users. Because there is no discernible indicator of increased users' influence, it appears that little could be gained by focusing more on users in the future either. Instead, junkets during which distributors invite professional critics to view a forthcoming film and interview its stars still represent valuable investments (Brunet et al. 2008).

For practitioners in other fields, who seek popular or recognized stars to support their marketing initiatives, our results also offer some insights. For example, for a new restaurant venue, would it be better to associate with a popular chef from a widely-distributed television show or a Michelin-starred chef who is not necessarily a household name? Our results suggest that the latter option is more effective, and the recent string of celebrity chef restaurant failures anecdotally supports this assertion (Perry and Hyde 2014). When stars associate with lower quality offerings, it also can erode the effectiveness of the celebrity endorsement. For example, the failed attempt by the social media company MySpace to revitalize its brand with Justin Timberlake's endorsement (Stockdale 2015), or the controversial endorsement of high-calorie soft drinks by One Direction and Taylor Swift (Sifferlin 2016), likely increased the variability of product quality endorsements overall, which ultimately can reduce their diagnosticity for consumers.

Expert opinions instead might matter more for hedonic and experiential products that cause consumers to confront high quality uncertainty. The impact of expert opinions on sales even may be stronger for products that require substantial expertise to be truly appreciated, such as wine, high-end restaurants, opera, or classical music. For example, losing or gaining a star in the Michelin guide can make or break a fine-dining restaurant, regardless of the popular opinions of non-expert diners (Fátharta 2016).

Limitations and further research

This meta-analytic investigation provides insights about the effects of star brand equity and product reviews for motion picture theatrical success. To support the dual purpose of addressing debates and challenging assumptions, we encourage further research along these lines. First, our

confirmation of the assumption of fading star power, as well as the preliminary evidence that movie quality variance explains this phenomenon, warrants further investigation. Delving deeper into the explanation for fading star power could give studio executives more concrete insights into how to slow down, or prevent, this decline. We also did not test the fading star power assumption in relation to risk. A recent study indicates that one of the main benefits of casting expensive stars is their ability to reduce movie revenue volatility in the first few weeks of the release (Joshi 2015). We thus encourage studies that investigate the capacity of stars to alleviate risk and further meta-analyses that challenge the fading star power assumption.

Second, we focus on star brand equity and product reviews, due to the prominence, accessibility, and diagnosticity of these cues (Basuroy et al. 2006; Joshi and Mao 2012). Yet other BO drivers, including sequels, franchises, film budget, or promotional expenses, also emit cues that may reduce consumers' quality uncertainty. We control for the effects of these and other focal cues in our analyses, but more specific statistical controls could be applied in further research. As studies of the drivers of BO performance continue to accumulate, eventually they should evolve to feature comprehensive models that all include the same cues. At that juncture, it would become possible to conduct a meta-analysis on any focal cue while controlling for the effects of the others by focusing on this pool of primary studies.

Third, it would be interesting to test whether our findings apply to motion pictures in other countries—and particularly China, considering the dramatic growth of its movie industry (i.e., 49% increase in BO revenue from 2015 to 2016; *The Economist* 2016). Our results hold for movies released in the United States, Europe, and an "other regions" category, but our principal aim for that test was statistical control. The effects are not necessarily generalizable across geographies and cultures. As Akdeniz and Talay (2013) note, country culture can modify the strength of the effects of star brand equity and product reviews. In addition, the congruence between a movie's cultural content and the national cultural background of its audience positively affects how it is received (Song et al. 2017). Thus, whether the relationships we identify hold in the cultural milieu of China or other fast-growing film markets worldwide is a key empirical question that further research should address.

In conclusion, we cite the screenwriter and author William Goldman, who summarized Hollywood in 1983 by noting, "Nobody knows anything"—a curt but accurate evaluation of

what the industry knew about how to create a box office hit (De Vany 2004). With our novel

approach, we advance ongoing discussions and resolve some persistent misconceptions about

two prominent, extensively scrutinized drivers of motion picture performance: star brand equity

and product reviews. By harnessing the collective power of extant empirical literature, we seek

to provide a more cohesive view of the field, while also preventing unsettled debates and

dogmatic assumptions from taking hold in either research or managerial circles.

Appendix

	Effect size (r)	Random sampling	Country US	Country other	Artistic star power	Media star power	Other perf. measures	Short-term BO	Median year	Range	Actor	Director	Other	Published	Top-tier	Zero-order	Movie awards	Study precision	Artistic dimension included	Market dimension included	Media dimension included	Critics' reviews included
Random sampling	.218**																					
Country US	-.001	.296**																				
Country other	-.049	-.238**	-.498**																			
Artistic star power	-.084	.135*	.113*	.018																		
Media star power	.099	-.084	-.249**	.165*	-.363**																	
Other perf. measures	-.092	.058	.050	-.038	.219**	-.214**																
Short-term BO	.015	-.158**	-.158**	.044	-.209**	.225**	-.357**															
Median year	-.155**	-.234**	-.286**	.192**	-.148*	.121*	-.243**	.227**														
Range	.059	.173**	.154**	-.128*	-.039	-.080	.032	-.093	-.209**													
Actor	.024	-.036	-.102	.080	.021	.168**	-.044	.114*	.087	-.023												
Director	.017	.246**	.128*	-.108	.258**	-.187**	.114*	-.203**	-.120*	.068	-.632**											
Other	-.071	.172**	.182**	-.105	.313**	-.197**	-.037	-.158**	-.141*	.197**	-.061	.175**										
Published	.036	-.135*	-.121*	.116*	-.105	.203**	.035	.034	-.080	-.172**	.036	-.061	-.245**									
Top-tier	-.079	-.231**	-.166**	.108	-.016	.147*	-.157**	.410**	.031	-.173**	.065	-.223**	-.055	.242**								
Zero-order	.421**	-.020	-.205**	.103	-.058	.244**	-.025	.313**	.032	-.051	.123*	-.133*	-.029	.114*	.304**							
Movie awards	-.056	.084	.105	-.069	.169**	-.009	.221**	-.234**	-.110	.072	.127*	.110	.082	.002	-.221**	-.079						
Study precision	.106	.296**	.186**	.077	.002	-.065	-.037	-.030	.150**	.492**	-.044	.025	.146*	-.219**	-.143*	.057	-.065					
Artistic dimension included	-.110	.054	.196**	-.111	-.117*	-.111	-.028	-.056	-.018	.012	-.096	.127*	.096	-.206**	-.042	-.204**	.145*	.075				
Market dimension included	-.130*	.045	.211**	-.105	.496**	-.151**	-.016	-.038	-.021	.001	-.057	.097	.112	-.181**	.005	-.192**	.064	.057	-.011			
Media dimension included	-.046	-.007	-.005	.059	.111	-.127*	-.075	.071	.017	-.006	.098	-.027	.072	.054	-.106	-.118*	.027	-.008	.048	.054		
Critics' reviews included	-.273**	.096	.173**	-.147*	.121*	-.222**	.134*	-.014	.107	-.030	.060	.134*	.032	-.192**	-.157**	-.525**	.172**	-.089	.245**	.157**	.177**	
Users' reviews included	-.172**	-.125*	.031	-.014	-.116*	-.094	-.021	-.058	.152**	-.023	.015	-.002	-.061	-.097	-.111	-.254**	.009	-.139*	-.044	-.037	-.051	.223**

Panel A: correlation matrix of variables in the star brand equity analyses

*: $p < .05$; **: $p < .01$.

Panel B: correlation matrix of variables in the product reviews analyses

	Effect size (r)	Random sampling	Country US	Country other	Users' reviews	Review volume	Other perf. measures	Short-term BO	Median year	Range	Published	Top-tier	Zero-order	Movie awards	Study precision	Artistic dimension included	Market dimension included	Media dimension included	Critics' reviews included

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Random sampling	.013																		
Country US	.084	.206**																	
Country other	-.128*	-.225**	-.574**																
Users' reviews	.082	.087	.057	.094															
Review volume	.248**	.170**	.114*	-.005	.227**														
Other perf. measures	.064	-.003	-.216**	.109*	.029	.003													
Short-term BO	-.221**	-.094	-.051	-.018	-.204**	-.114*	-.347**												
Median year	-.016	.062	-.278**	.257**	.279**	.116*	-.095	.083											
Range	-.096	-.092	.036	-.085	-.163**	-.216**	.008	.033	-.545**										
Published	-.027	-.031	-.242**	.139*	-.055	-.009	.109*	.159**	-.044	-.010									
Top-tier	-.253**	-.035	-.060	.026	-.109*	-.089	-.204**	.413**	.051	-.075	.213**								
Zero-order	.035	-.144**	-.015	-.042	-.134*	-.105	-.048	.223**	.094	.062	.184**	.125*							
Movie awards	-.014	-.014	.085	-.034	-.126*	-.030	.093	-.152**	-.393**	.096	-.116*	-.193**	-.089						
Study precision	-.142**	.070	.004	-.101	-.206**	-.246**	.014	.151**	-.122*	.493**	-.156**	-.112*	.121*	.018					
Artistic dimension included	-.046	.110*	.156**	-.117*	-.212**	.019	.122*	-.108*	-.220**	.003	-.175**	-.104	-.206**	.207**	.046				
Market dimension included	-.097	-.025	-.056	.013	-.091	-.058	.110*	-.031	-.126*	.062	-.096	-.053	-.312**	.049	.037	.189**			
Media dimension included	-.005	.119*	.129*	-.108*	-.001	-.040	-.143**	.020	-.042	.015	.034	.186**	-.281**	-.012	.037	.005	-.137*		
Critics' reviews included	-.093	.132*	.048	-.011	.572**	.152**	-.030	-.117*	.163**	-.142**	-.121*	-.016	-.312**	-.062	-.084	-.105	-.010	.186**	
Users' reviews included	-.058	-.005	-.027	-.038	-.329**	.012	-.001	-.028	.108*	-.109*	-.095	-.058	-.288**	-.046	-.025	-.093	.039	.123*	-.188**

*: $p < .05$; **: $p < .01$.

References

Adler, M. (1985). Stardom and talent. *American Economic Review, 75*(1), 208-212.

Akdeniz, M. B., Calantone, R. J., & Voorhees, C. M. (2013). Effectiveness of marketing cues on consumer perceptions of quality: The moderating roles of brand reputation and third-party information. *Psychology & Marketing*, *30*(1), 76-89.

Akdeniz, M. B. & Talay, M. B. (2013). Cultural variations in the use of marketing signals: A multilevel analysis of the motion picture industry. *Journal of the Academy of Marketing Science, 41*(5), 601-624.

Akerlof, G. A. (1970). The market for 'lemons': quality uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*(3), 488-500.

Albert, S. (1998). Movie stars and the distribution of financially successful films in the motion picture industry. *Journal of Cultural Economics, 22*(4), 249-270.

Anand, N., & Watson, M. R. (2004). Tournament rituals in the evolution of fields: the case of the Grammy awards. *Academy of Management Journal*, 47(1), 59-80.

Babić Rosario, A. Sotgiu, F., Valck, K. D., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research, 53*(3), 297-318.

Bagella, M., & Becchetti, L. (1999). The determinants of motion picture box office performance: Evidence from movies produced in Italy. *Journal of Cultural Economics, 23*(4), 237-256.

Balio, T. (1985). *The American film industry.* Madison: University of Wisconsin Press.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing, 67*(4), 103-117.

Basuroy, S., Desai, K. K., & Talukdar, D. (2006). An empirical investigation of signaling in the motion picture industry. *Journal of Marketing Research, 43*(2), 287-295.

Bijmolt, T. H. A. & Pieters, R. G. M. (2001). Meta-analysis in marketing when studies contain multiple measurements. *Marketing Letters, 12*(2), 157-169.

Boatwright, P., Basuroy, S., & Kamakura, W. (2007). Reviewing the reviewers: The impact of individual film critics on box-office performance. *Quantitative Marketing and Economics*, *5*(4), 401-425.

Borenstein, M. (1994). Effect size for continuous data. In *The Handbook of Research Synthesis and Meta-Analysis*, H. Cooper, L.V. Hedges, and J. Valentine, Eds. New York: Russel Edge Foundation.

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.

Brightlocal (2016). Local consumer review survey. Retrieved on June 19, 2017, from: https://www.brightlocal.com/learn/local-consumer-review-survey

Brunet, J., Legoux, R., & Najar, S. (2008). Controlled and uncontrolled buzz as a marketing strategy in the film industry: the effectiveness of the junkets. Paper presented at the meeting *Cultural Production in a Global Context: The Worldwide Film Industries*, Copenhagen, Denmark.

Bughin, J., Doogan, J., & Vetvik, O. J. (2010). A new way to measure word-of-mouth marketing. *McKinsey Quarterly, 2*, 113-116.

Chakravarty, A., Liu, Y., & Mazumdar, T. (2010). The differential effects of online word-of-mouth and critics' reviews on pre-release movie evaluation. *Journal of Interactive Marketing*, *24*(3), 185-197.

Chang, B.-H., & Ki, E.-J. (2005). Devising a practical model for predicting theatrical movie success: focusing on the experience good property. *Journal of Media Economics*, *18*(4), 247-269.

Chen, Y., Liu, Y., & Zhang, J. (2012). When do third-party product reviews affect firm value and what can firms do? The case of media critics and professional movie reviews. *Journal of Marketing, 76*(2), 116-134.

Chisholm, D. C., Fernandez-Blanco, V., Ravid, S. A., & Walls, W. D. (2015). Economics of motion pictures: The state of the art. *Journal of Cultural Economics, 39*(1), 1-13.

d'Astous, A., & Touil, N. (1999). Consumer evaluations of movies on the basis of critics' judgments. *Psychology & Marketing, 16*(8), 677-694.

De Matos, C. A., & Rossi, C. A. V. (2008). Word-of-mouth communications in marketing: a meta-analytic review of the antecedents and moderators. *Journal of the Academy of Marketing Science, 36*(4), 578-596.

Desai, K. K., & Basuroy, S. (2005). Interactive influence of genre familiarity, star power, and critics' reviews in the cultural goods industry: The case of motion pictures. *Psychology & Marketing*, *22*(3), 203-223.

Deuchert, E., Adjamah, K., & Florian, P. (2005). For Oscar glory or Oscar money? *Journal of Cultural Economics*, *29*(3), 159–176.

De Vany, A. (2004). *Hollywood economics: How extreme uncertainty shapes the film industry*. New York: Routledge.

Duan, W., Gu, B., & Whinston, A. B. (2008). The dynamics of online word-of-mouth and product sales-an empirical investigation of the movie industry. *Journal of Retailing, 84*(2), 233-242.

Duval S., & Tweedie R. (2000). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89-98.

Eisend, M., & Tarrahi, F. (2014). Meta-analysis selection bias in marketing research. *International Journal of Research in Marketing*, *31*(3), 317-326.

Elberse, A. (2007). The power of stars: Do star actors drive the success of movies? *Journal of Marketing, 71*(4), 102-120.

Eliashberg, J., & Sawhney, M. S. (1994). Modeling goes to Hollywood: Predicting individual differences in movie enjoyment. *Management Science*, *40*(9), 1151-1173.

Eliashberg, J. & Shugan, S. M. (1997). Film critics: Influencers or predictors? *Journal of Marketing, 61*(April), 68-78.

Fátharta, C. Ó. (2016). Kevin Thornton's restaurant to close after loss of Michelin star. *Irish Examiner*. Retrieved on April 10, 2017 from http://www.irishexaminer.com/ireland/kevin-thorntons-restaurant-to-close-after-loss-of-michelin-star-418956.html

Feldman, J. M., & Lynch, J. G. (1988). Self-generated validity and other effects of measurement on belief, attitude, intention, and behavior. *Journal of Applied Psychology*, *73*(3), 421-435.

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods*, *17*(1), 120.

Flanagin, A. J. & Metzger, M. J. (2013). Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics. *Computers in Human Behavior, 29*(4), 1626-1634.

Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing*, *90*(2), 217-232.

Gazley, A., Clark, G., & Sinha, A. (2011). Understanding preferences for motion pictures. *Journal of Business Research, 64*(8), 854-861.

Hadida, A. L. (2009). Motion picture performance: A review and research agenda. *International Journal of Management Reviews, 11*(3), 297-335.

Hadida, A. L. (2010). Commercial success and artistic recognition of motion picture projects. *Journal of Cultural Economics*, *34*(1), 45-80.

Hennig-Thurau, T., Walsh, G., & Wruck, O. (2001). An investigation of the factors determining the success of service innovations: the case of motion pictures. *Academy of Marketing Science Review*, *2001*(6), 1-23.

Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science, 43*(3), 375-394.

Hofmann, J., Clement, M., Völckner, F., & Hennig-Thurau, T. (2017). Empirical generalizations on the impact of stars on the economic success of movies. *International Journal of Research in Marketing*, *34*(2), 442-461.

Holbrook, M. B., & Addis, M. (2007). Taste versus the market: an extension of research on the consumption of popular culture. *Journal of Consumer Research, 34*(October), 415-424.

Hunter, J. E. & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings.* Thousand Oaks, CA: Sage Publications.

Internet Movie Database (2016). Advanced title search. Available at: http://www.imdb.com/search/title.

Johnson, B. T. & Huedo-Medina, T. B. (2011). Depicting estimates using the intercept in meta-regression models: The moving constant technique. *Research synthesis methods, 2*(3), 204-220.

Johnston, N. (2009). *IMDb movie ratings over the years*. Retrieved on March 11, 2017 from: http://www.njohnston.ca/2009/10/imdb-movie-ratings-over-the-years/

Joshi, A. (2015). Movie stars and the volatility of movie revenues. *Journal of Media Economics, 28*(4), 246-267.

Joshi, A. M. & Hanssens, D. M. (2009). Movie advertising and the stock market valuation of studios: A case of "great expectations? *Marketing Science, 28*(2), 239-250.

Joshi, A., & Mao, H. (2012). Adapting to succeed? Leveraging the brand equity of best sellers to succeed at the box office. *Journal of the Academy of Marketing Science*, *40*(4), 558-571.

Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods*, *15(*4), 624-662.

Kermode, M. (2014). *Hatchet job: Love movies, hate critics.* London, UK: Pan MacMillan.

Knapp, A. K., Hennig-Thurau, T., & Mathys, J. (2014). The importance of reciprocal spillover effects for the valuation of bestseller brands: introducing and testing a contingency model. *Journal of the Academy of Marketing Science*, *42*(2), 205-221.

Knoll, J., & Matthes, J. (2017). The effectiveness of celebrity endorsements: a meta-analysis. *Journal of the Academy of Marketing Science*, *45*(1), 55-75.

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2*(1), 61-76.

Langan, R., Besharat, A., & Varki, S. (2017). The effect of review valence and variance on product evaluations: An examination of intrinsic and extrinsic cues. *International Journal of Research in Marketing, 34*(2), 414-429.

Legoux, R., Larocque, D., Laporte, S., Belmati, S., & Boquet, T. (2016). The effect of critical reviews on exhibitors' decisions: Do reviews affect the survival of a movie on screen? *International Journal of Research in Marketing*, *33*(2), 357-374

Levin, A. M., Levin, I. P., & Heath, C. E. (1997). Movie stars and authors as brand names: measuring brand equity in experiential products. *Advances in Consumer Research, 24*(1), 175-181.

Lin, D.-Y. & Sullivan, P. F. (2009). Meta-analysis of genome-wide association studies with overlapping subjects. *American Journal of Human Genetics, 85*(December), 862-872.

Lipsey, M. W. & Wilson, D. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage Publications.

Litman, B. R. (1983). Predicting success of theatrical movies: an empirical study. *Journal of Popular Culture, 16*(4), 156-175.

Liu, A., Liu, Y., & Mazumdar, T. (2014). Star power in the eye of the beholder: A study of the influence of stars in the movie industry. *Marketing Letters, 25*(4), 385-396.

Liu, Y. (2006). Word of mouth for movies: its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74-89.

Luo, L., Chen, X., Han, J., & Whan Park, C. (2010). Dilution and enhancement of celebrity brands through sequential movie releases. *Journal of Marketing Research, 47*(6), 1114-1128.

Lynch, J. G., Marmorstein, H., & Weigold, M. F. (1988). Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. *Journal of Consumer Research*, *15*(2), 169-184.

Marich, R. (2014). Movie star power dims as actors become hucksters. *Huffington Post*. Retrieved on March, 27, 2017 from http://www.huffingtonpost.com/robert-marich/movie-star-power-dims-as-_b_5556322.html

McClintock, P. & Galloway, S. (2017). Matt Damon's 'The Great Wall' to lose $75 million; future U.S.-China productions in doubt. *Hollywood Reporter*. Retrieved on April 2, 2017 from http://www.hollywoodreporter.com/news/what-great-walls-box-office-flop-will-cost-studios-981602

Merry, S. (2015). Are A-list actors a waste of money? *Sydney Morning Herald*. Retrieved on August 7, 2016 from http://www.smh.com.au/entertainment/movies/are-alist-actors-a-waste-of-money-20151125-gl7f0i.html

Mukhopadhyay, P. (2008). *Theory and methods of survey sampling*. New Delhi: PHI Learning Pvt. Ltd.

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: an overview of reviews. *Clinical Psychology Review*, *33*(4), 501-511.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, *78*(2), 311-329.

Nielsen (2015). Global trust in advertising. Retrieved on August 7, 2016 from: http://www.nielsen.com/us/en/insights/reports/2015/global-trust-in-advertising-2015.html].

Patten, D. (2017). Oscar ratings down double digits in demo; 32.9m viewership hits 9-year low. *Deadline Hollywood*. Retrieved on March 31, 2017 from http://deadline.com/2017/02/oscars-ratings-2017-down-best-picture-mistake-jimmy-kimmel-abc-1202009034/

Perry, K. and Hyde, D. (2014). Celebrity chefs' restaurants among most overpriced and worst food. *The Telegraph*. Retrieved on April 10, 2017 from http://www.telegraph.co.uk/foodanddrink/11272075/Celebrity-chefs-restaurants-among-most-overpriced-and-worst-food.html

Politis, D. N., Romano, J. P., & Wolf, M. (2001). On the asymptotic theory of subsampling. *Statistica Sinica*, *11*(4), 1105-1124.

Proserpio, D., & Zervas, G. (2017). Online reputation management: Estimating the impact of management responses on consumer reviews. *Marketing Science*, https://doi.org/10.1287/mksc.2017.1043, 1-21.

Purohit, D., & Srivastava, J. (2001). Effect of manufacturer reputation, retailer reputation, and product warranty on consumer judgments of product quality: A cue diagnosticity framework. *Journal of Consumer Psychology*, *10*(3), 123-134.

PwC (2015). Box office resilience underscores the continuing popularity of the cinematic experience. Retrieved on August 7, 2016 from: http://www.pwc.com/gx/en/global-entertainment-media-outlook/segment-insights/filmed-entertainment.jhtml].

Ravid, S. A. (1999). Information, blockbusters, and stars: A study of the film industry. *Journal of Business Research, 72*(4), 463-492.

Reddy, S. K., Swaminathan, V., & Motley, C. M. (1998). Exploring the determinants of Broadway show success. *Journal of Marketing Research, 35*(3), 370-383.

Reinstein, D. A. & Snyder, C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *Journal of Industrial Economics, 53*(1), 27-51.

Rosen, S. (1981). The economics of superstars. *American Economic Review, 71*(5), 845-858.

Rosenberg, M. S. (2005). The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. *Evolution*, *59*(2), 464-468.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, *86*(3), 638-641.

Sedgwick, J. (1999). The comparative popularity of stars in mid-1930s Britain. *Journal of Popular British Cinema*, *2*, 121-128.

Sifferlin, A. (2016). Food endorsed by celebrities are almost all terrible for you. Retrieved on June 8, 2017 from: http://time.com/4358156/celebrity-endorsements-food-soda-drink/

Song, R., Moon, S., Chen, H. A., & Houston, M. B. (2017). When marketing strategy meets culture: The role of culture in product evaluations. *Journal of the Academy of Marketing Science*, DOI 10.1007/s11747-017-0525-x 1-19.

Sood, S. & Drèze, X. (2006). Brand extensions of experiential goods: Movie sequel evaluations. *Journal of Consumer Research, 33*(3), 352-360.

Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business* (Vol. 5). New York: Routledge.

Sterne, J. A. & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. In A. R. Rothstein, A. J. Sutton, & M. Borenstein (eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments.* New York: Wiley.

Stigler, G. J., & Becker, G. S. (1977). De gustibus non est disputandum. *American Economic Review*, *67*(2), 76-90.

Stockdale, D. (2015). 10 worst products endorsed by celebrities. Retrieved on June 8, 2017 from: http://www.twcc.com/articles/2015/12/28/1/10-worst-products-endorsed-by-celebrities.3

Suárez-Vázquez, A. (2011). Critic power or star power? The influence of hallmarks of quality of motion pictures: An experimental approach. *Journal of Cultural Economics, 35*(2), 119-135.

The Economist (2016). Fading stars: Hollywood studios can no longer bank on the pulling-power of famous actors. Retrieved on August 8, 2016 from: http://www.economist.com/news/business/21693591-hollywood-studios-can-no-longer-bank-pulling-power-famous-actors-fading-stars].

Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: social contagion versus marketing effort. *American Journal of Sociology*, *106*(5), 1409-1435.

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*(4), 1274-1294.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the Metafor package. *Journal of Statistical Software, 63*(3), 1-48.

Watson, IV, G. F., Beck, J. T., Henderson, C. M., & Palmatier, R. W. (2015). Building, measuring, and profiting from customer loyalty. *Journal of the Academy of Marketing Science, 43*(6), 790-825.

You, Y., Vadakkepatt, G. G., & Joshi, A. (2015). A meta-analysis of electronic word-of-mouth elasticity. *Journal of Marketing, 79*(2), 19-39.

Zhang, X. & Dellarocas, C. (2006). The lord of the ratings: Is a movie's fate is influenced by reviews? Proceedings of the *International Conference on Information Systems*, Milwaukee, WI.

Zuckerman, E. W., & Kim, T. Y. (2003). The critical trade-off: Identity assignment and box-office success in the feature film industry. *Industrial and Corporate Change*, *12*(1), 27-67.

Zuckerman, E. W., Kim, T. Y., Ukanwa, K., & Rittmann, J. V. (2003). Robust identities or nonentities? Typecasting in the feature-film labor market. *American Journal of Sociology, 108*(5), 1018-1073.

Table 1 Comparison between the current meta-analysis and previous meta-analyses on star brand equity and product reviews

Study[a]	Focus	Primary studies[b]	Star contributors	Product reviews	Effect size	Sales variable(s)	Meta-analytic model	Publication bias	Subject overlap in primary studies
Hofmann et al. (2017)	Star brand equity	61	Actors	-	r	Short-term and Long-term BO revenue	Weighted random-effect	Fail safe N	Not taken into account
Floyd et al. (2014)	Product reviews	4	-	Users and critics	Elasticities	BO revenue	Hierarchical weighted random effect	Manuscripts status; call for unpublished studies	Not taken into account
You et al. (2015)	Product reviews	14	-	Users	Elasticities	BO revenue	Hierarchical weighted random effect	Manuscripts status	Not taken into account
Babić Rosario et al. (2016)	Product reviews	26	-	Users	r	BO revenue	Hierarchical weighted random effect	Fail safe N; manuscripts status; call for unpublished studies; effect size precision	Not taken into account
Current study	Star brand equity	110	Actors, directors, and others		r	Short-term and Long-term BO revenue; Other theatrical performance indicators	Hierarchical weighted random effect	Fail safe N; manuscripts status; call for unpublished studies; effect size precision	Adjustment of primary study sample sizes and drawing of non-overlapping sub-samples
	Product reviews	112		Users and critics					

a: de Matos and Rossi's (2008) meta-analysis focuses on the antecedents of users' reviews unlike the studies in this table that focus on their consequences.

b: In Babić Rosario et al. (2016) as well as Floyd et al. (2014) and You et al. (2015), this refers to the number of studies about movies only as these meta-analyses investigate multiple product categories and have a higher total number of studies.

Table 2 Coding scheme of variables in the meta-analysis

Variable	Coding scheme
Dependent variable	
Movie theatrical performance	First dummy: 1 if effect size based on gross domestic BO revenue garnered by a given movie during its opening week-end or week and 0 otherwise. Second dummy: 1 if effect size based on other indicators of BO performance (theatrical admissions, international BO, rentals, return on investment, length of theatrical run, and sales per screen) and 0 otherwise. Reference: long-term BO (i.e., gross domestic BO revenue obtained during the entire theatrical run).
Star power	
Brand equity dimensions	First dummy: 1 if effect size based on whether a movie contributor has ever received or been nominated for an award or if effect size based on the number of awards received or been nominated for (artistic dimension) and 0 otherwise. Second dummy: 1 if effect size based on the inclusion of stars on media lists and 0 otherwise. Reference: market dimension (operationalized as the BO of the movies stars contributed to in the past: number of appearances in top grossing movies, cumulative BO earnings over career or in recent years).
Contributors	First dummy: 1 if stars include actors and 0 otherwise. Second dummy: 1 if stars include directors and 0 otherwise. Third dummy: 1 if stars include "other" contributors (e.g., composers) and zero otherwise.
Product reviews	
Source	Dummy: 1 if the reviews come from users (regular consumers) and 0 if they come from professional critics.
Metric	Dummy: 1 if the volume of reviews (i.e., the number of reviews available at a particular time) was used and 0 if the valence of reviews (i.e., numerical summary evaluation) was used.
Time	
Movies' median year of release	Continuous: operationalized as the midpoint between start and end of the observation window of the study (for instance, 1984 would be the "median year" of an effect size from a primary study sampling movies from 1956 to 2012). Mean-centered.
Movies' range of release years	Continuous: difference between the earliest and latest year of movie release. Mean-centered.
Study characteristics	
Sampling	Dummy: 1 when movies were sampled randomly (e.g., 200 movies selected from the past 5 years) and 0 when they were drawn through convenience sampling (e.g., top 100 grossing movies in a given period or movies chosen for specific characteristics).
Country	First dummy: 1 if the BO data were collected in the US and 0 otherwise. Second dummy: 1 if the BO data were collected in "other" countries (e.g., Southeast Asia, South America, Australia) and zero otherwise. Reference: Europe.
Status	Dummy: 1 when the study was published and 0 otherwise (i.e., working papers, dissertations).
Movie awards	Dummy: 1 when the study controlled for whether the sampled movies received or were nominated for an industry award (e.g., Golden Globes, Oscars).
Top-tier journal	Dummy: 1 when the study was published in one of the *Financial Times* list of 50 top business journals and 0 otherwise.[10]
Zero-order correlation	Dummy: 1 when the effect size came from a correlation matrix and 0 when it was converted to a correlation from the statistical test of a model parameter.
Effect size precision	Continuous: inverse of the standard error of the Fisher z-transformed correlations. Mean-centered.
Inclusion of other focal effect sizes	Five dummies: 1 when effect size is from a model in which the effect of the artistic dimension, the market dimension, the media dimension, critics' reviews or users' reviews are also included and zero otherwise, respectively.

---

[10] The *American Journal of Sociology* was considered a top journal due to its high impact factor.

Bivariate relationship		# effect sizes (k)	Total N[a]	Mean effect size	Confidence interval		Prediction interval		Effect size distribution heterogeneity			Publication bias analysis			
					Lower-Bound	Upper-Bound	Lower-Bound	Upper-Bound	Tau^2	I^2	Q-statistic	Egger's regression (t value)	Rosenthal's fail safe N	Rosenberg's Fail safe N	Trim-and-fill adjusted mean effect size[b]
Artistic star brand equity	Short	06	6,573	.084	-.007	.175	-.119	.281	.009	81%	35.4***	-.47	-	-	+.007
	Long	24	16,615	.123	.086	.159	-.016	.257	.005	66%	78.8***	1.33	1,237	902	-.026
	Other	19	11,396	.061	-.031	.152	-.211	.325	.017	83%	215.7**	-1.22	-	-	+.030
Market star brand equity	Short	41	29,045	.163	.087	.237	-.181	.471	.030	95%	893.2***	-4.91***	10,097	27,094	+.067
	Long	62	33,699	.146	.097	.195	-.155	.522	.023	94%	1276.1***	-.93	19,845	20,055	+ .037
	Other	28	12,018	.064	.019	.108	-.097	.221	.006	66%	84.8***	.087	415	281	-
Media star brand equity	Short	59	32,006	.114	.050	.177	-.126	.341	.014	91%	307.3***	4.53***	1,674	1,557	-.023
	Long	53	26,736	.189	.114	.263	-.217	.540	.043	94%	1173.5***	1.08	8,790	10,448	-
	Other	11	13,475	.176	.058	.289	-.179	.491	.030	97%	149.63***	-.06	120	116	+.047
Critics (valence)	Short	71	37,783	.153	.099	.206	-.154	.433	.024	90%	669.1***	1.05	10,719	10,997	-
	Long	67	26,130	.208	.175	.241	-.029	.423	.015	81%	417.0***	1.84	19,131	14,013	-.051
	Other	34	15,454	.268	.223	.312	-.022	.516	.022	90%	309.9***	-.12	9,572	9,071	+.021
Critics (volume)	Short	10	3,382	.282	.110	.438	-.269	.694	.075	95%	84.0***	1.27	489	393	-
	Long	16	4,182	.327	.221	.426	-.101	.653	.047	92%	127.1***	.88	1,909	1,561	-
	Other	6	1,595	.236	.099	.365	-.057	.491	.018	78%	19.3**	-.26	138	109	+.026
Users (valence)	Short	14	7,708	.102	.049	.155	-.047	.247	.005	31%	44.2***	.63	234	201	-.025
	Long	46	13,662	.202	.163	.240	.001	.386	.011	72%	219.5***	-1.04	8,688	7,437	+.028
	Other	18	7,016	.178	.097	.258	-.125	.451	.022	81%	118.0***	.92	784	493	-
Users (volume)	Short	8	1,021	.283	.084	.461	-.258	.689	.069	88%	47.6***	2.47*	149	78	-.046
	Long	27	4,360	.391	.237	.525	-.340	.827	.145	87%	508.4***	-.58	3,594	3,713	+.090
	Other	10	2,613	.285	.097	.454	-.315	.723	.091	92%	96.5***	-.69	422	524	+.046

Table 3 Results for the summary meta-analyses

*: $p < .05$, **: $p < .01$, ***: $p < .001$.

a: Based on the total number of movies used in each primary study; it does not reflect the adjusted sample sizes computed to account for movie overlap.

b: Based on the difference between the mean effect size obtained when using the trim-and-fill procedure to correct for asymmetry and the observed mean effect size, which suggests the possibility of publication bias. A nil adjustment indicates a symmetric funnel plot.

Table 4 Meta-regressions explanatory power

Bivariate relationship		Summary meta-analysis Tau^2 (1)	Total Tau^2 (2)	Meta-regression[a] Tau^2 accounted for (3)	Residual Tau^2 (4)	Between-effect sizes variance accounted for by meta-regression (3)/(2) x100
Artistic star brand equity	Short	.009				
	Long	.005				
	Other	.017				
Market star brand equity	Short	.030	.177	.166	.011	94%
	Long	.023				
	Other	.006				
Media star brand equity	Short	.014				
	Long	.043				
	Other	.030				
Critics (valence)	Short	.024				
	Long	.015				
	Other	.022				
Critics (volume)	Short	.075				
	Long	.047				
	Other	.018	.544	.515	.029	94%
Users (valence)	Short	.005				
	Long	.011				
	Other	.022				
Users (volume)	Short	.069				
	Long	.145				
	Other	.091				

a: Tau^2 and Residual Tau^2 from meta-regressions are based on Model 2 for star brand equity and Model 4 for product reviews, respectively.

Table 5 Influence of moderators on the relationship between movie theatrical performance and star brand equity

Predictors[a]	Model 1: Star brand equity effect sizes (k = 303) Par. estimate (SE)	Model 2: Star brand equity effect sizes with interactions terms (k = 303) Par. estimate (SE)
Intercept	.1509 (.0482)**	.1347 (.0467)**
Random sampling (k = 93)	.0443 (.0248)‡	.0475 (.0240)*
Other market (k = 45)	-.0141 (.0218)	-.0166 (.0219)
US market (k = 178)	-.0048 (.0229)	-.0029 (.0224)
Artistic dimension of star's brand equity (k = 49)	-.0397 (.0375)	-.0077 (.0432)
Media dimension of star's brand equity (k = 123)	-.0065 (.0247)	-.0013 (.0291)
Other performance measures (k = 58)	-.0930 (.0230)***	-.0973 (.0314)**
Short-term BO (k = 106)	-.0269 (.0187)	-.0130 (.0229)
Median year of movie release[b,c]	-.0021 (.0010)*	-.0032 (.0012)**
Range of movie release years[c]	-.0005 (.0018)	-.0007 (.0017)
Actor (k = 216)	-.0192 (.0249)	-.0219 (.0250)
Director (k = 112	-.0067 (.0242)	-.0078 (.0243)
Other contributor (k = 39)	-.0731 (.0237)***	-.0592 (.0245)*
Movie awards (k = 28)	-.0032 (.0365)	-.0234 (.0353)
Published (k = 270)	.0051 (.0359)	.00067 (.0344)
Top-tier publication (k = 98)	-.0611 (.0263)*	-.0569 (.0255)*
Zero-order correlation (k = 112)	.1716 (.0295)***	.1836 (.0281)***
Effect size precision[c]	.0007 (.0011)	.0006 (.0011)
Artistic dimension included (k = 20)	-.0368 (.0383)	-.0307 (.0367)
Market dimension included (k = 18)	-.0219 (.0446)	-.0269 (.0436)
Media dimension included (k = 7)	.0223 (.0534)	.0258 (.0523)
Critics' reviews included (k = 97)	.0101 (.0288)	.0124 (.0275)
Users' reviews included (k = 30)	-.0295 (.0374)	-.0137 (.0359)
Artistic dim. of star's brand equity X Other perf. measures	-	.0186 (.0531)
Media dim. of star's brand equity X Other perf. measures	-	.0213 (.0498)
Artistic dim. of star's brand equity X Short-term BO	-	-.0891 (.0485)‡
Media dim. of star's brand equity X Short-term BO	-	-.0141 (.0399)
Artistic dim. of star's brand equity X Median year[c]	-	.0051 (.0018)**
Media dim. of star's brand equity X Median year[c]	-	-.0008 (.0020)

a: For all the meta-regressions the reference category is market dimension, non-random sampling, European market, long-term BO, no control for movie awards, non-published, non top-tier journal, and partial correlation. The number of effect sizes (k) is indicated for the category displayed (e.g., there were 93 effect sizes obtained from random sampling and 303-93 = 210 from non-random sampling).
b: The reference year is 1998.
c: Mean-centered continuous predictor.
‡: $p < .1$, *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Table 6 Influence of moderators on the relationship between movie theatrical performance and product reviews

Predictors[a]	Model 3: Critics only[b] (k = 204) Par. estimate (SE)	Model 4: Critics + Users[c] (k = 331) Par. estimate (SE)
Intercept	.2292 (.0680)***	.3183 (.0643)***
Random sampling (k =57; k = 103)	.0203 (.0340)	-.0095 (.0310)
Other market (k = 22; k = 44)	-.0098 (.0350)	-.0529 (.0374)
US market (k = 135; k = 226)	.0355 (.0388)	.0068 (.0376)
Users' review (k = 0; k = 127)	-	-.0270 (.0338)
Review volume (k = 32; k = 77)	.1282 (.0382)***	.0841 (.0343)*
Other performance measures (k = 40; k = 68)	.0414 (.0357)	-.0144 (.0306)
Short-term BO (k = 81; k = 107)	-.0526 (.0273)‡	-.0656 (.0246)**
Median year of movie release[d]	-.0038 (.0029)	-.0027 (.0031)
Range of movie release years[d]	-.0014 (.0023)	-.0030 (.0025)
Movie awards (k = 25; k = 31)	-.0470 (.0453)	-.0650 (.0449)
Published (k = 184; k = 294)	-.0622 (.0504)	-.0241 (.0447)
Top-tier publication (k = 62; k = 88)	-.0897 (.0429)	-.1375 (.0393)***
Zero-order correlation (k = 76; k = 107)	.0563 (.0400)	.0267 (.0378)
Effect size precision[d]	-.0015 (.0017)	-.0027 (.0017)
Artistic dimension included (k =26; k = 27)	-.0470 (.0453)	-.0775 (.0488)
Market dimension included (k = 56; k = 40)	-.0133 (.0413)	-.0185 (.0383)
Media dimension included (k = 47; k = 29)	.0433 (.0426)	.0385 (.0413)
Critics dimension included (k = 0; k = 56)	-	-.1687 (.0440)
Users' reviews included (k = 49; k = 71)	.0162 (.0240)	-.1131 (.0425)
Users' review X Review volume	-	.0937 (.0480)‡
Users' review X Median year of movie release[d]	-	-.0026 (.0055)
Other performance measures X Review volume	-.1169 (.0682)‡	-
Short-term BO X Review volume	-.0156 (.0585)	-

a: For all the meta-regressions the reference category is non-random sampling, other market, critics' reviews, review valence, long-term BO, non-published, non top-tier journal, partial correlation. The number of effect sizes (k) is indicated for the category displayed for Model 3 and Model 4, respectively (e.g., Model 3: there were 57 effect sizes obtained from random sampling and 204-57 = 147 from non-random sampling. Model 4: there were 103 effect sizes obtained from random sampling and 331-103 = 228 from non-random sampling).
b: The reference year for median year is 2001.
c: The reference year for median year is 2003.
d: Mean-centered continuous predictor.
‡: $p < .1$, *: $p < .05$, **: $p < .01$, ***: $p < .001$.

Figure 1 Debates and Assumptions

Star brand equity

Debate

1. Popularity versus artistic recognition. Does star popularity (market and media dimensions) or artistic recognition (artistic dimension) have a stronger impact on BO performance?

Assumption

2. Fading star power. The impact of star brand equity on movie performance is weakening over time.


Product reviews

Debate

3. Influencer versus predictor effect. What is the impact of critics' reviews on short-term BO compared to long-term BO?

Assumptions

4. Users' advantage. Users' reviews are a stronger driver of BO performance than professional critics' reviews.

5. Increasing users' advantage. The impact of users' reviews on BO performance is strengthening over time whereas the impact of critics' reviews is weakening.
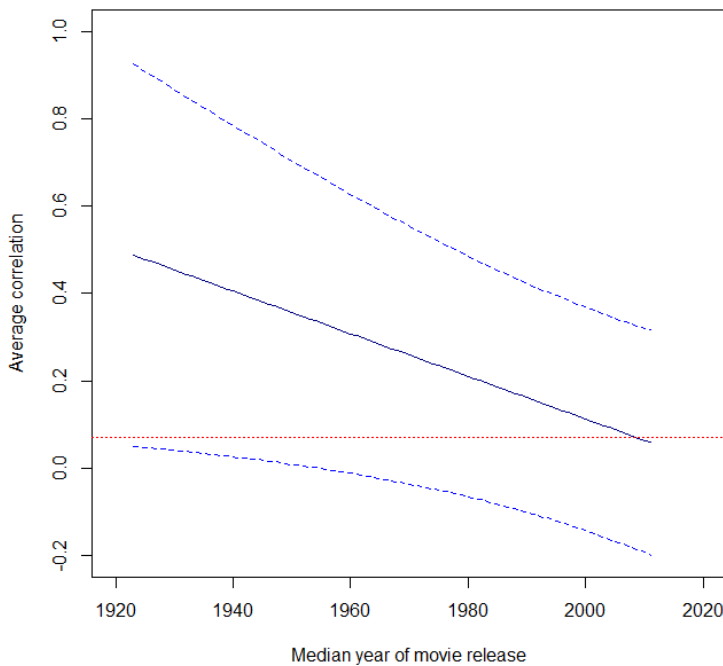

Fig. 2a Evolution of stars' market brand equity-movie theatrical performance mean effect size over time with 95% confidence bands
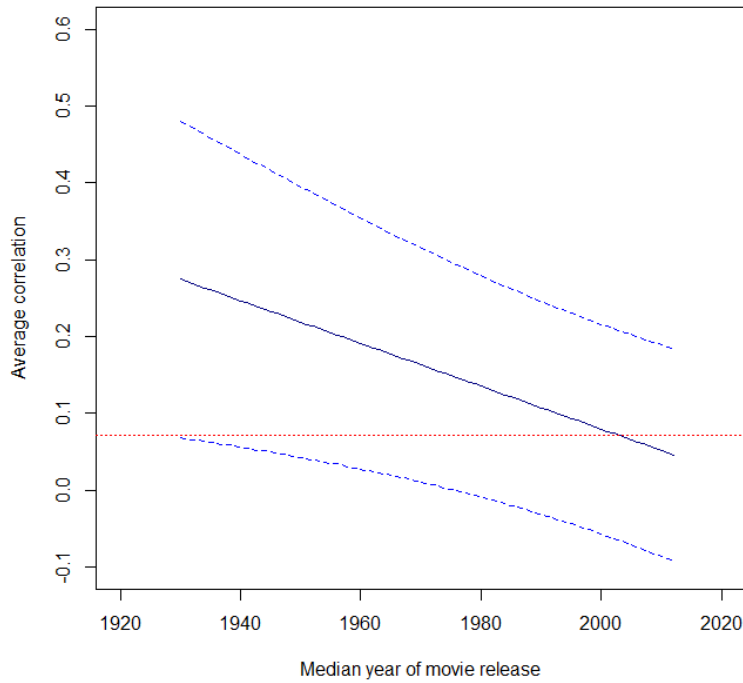
Meta-regression equation:

y = .113 -.003 (median year of movie release) +.001 (range of movie years) -.095 (other measures) -.011 (short-term BO) -.072 (other contributors) + .182 (zero order correlation)

Fig. 2b Evolution of stars' media brand equity-movie theatrical performance mean effect size over time with 95 % confidence bands

Meta-regression equation:

y = .135 -.005 (median year of movie release) +.001 (range of movie years) -.111 (other measures) -.044 (short-term BO) -.120 (other contributors) + .211 (zero order correlation)

·········· Mean effect size for the artistic dimension of star brand equity over the period ($\bar{r}$ = .072; 95%CI: .029 - .116).

Fig. 3a Forest plot of stars' market brand equity- movie theatrical performance effect sizes sorted from 1935 to 1999
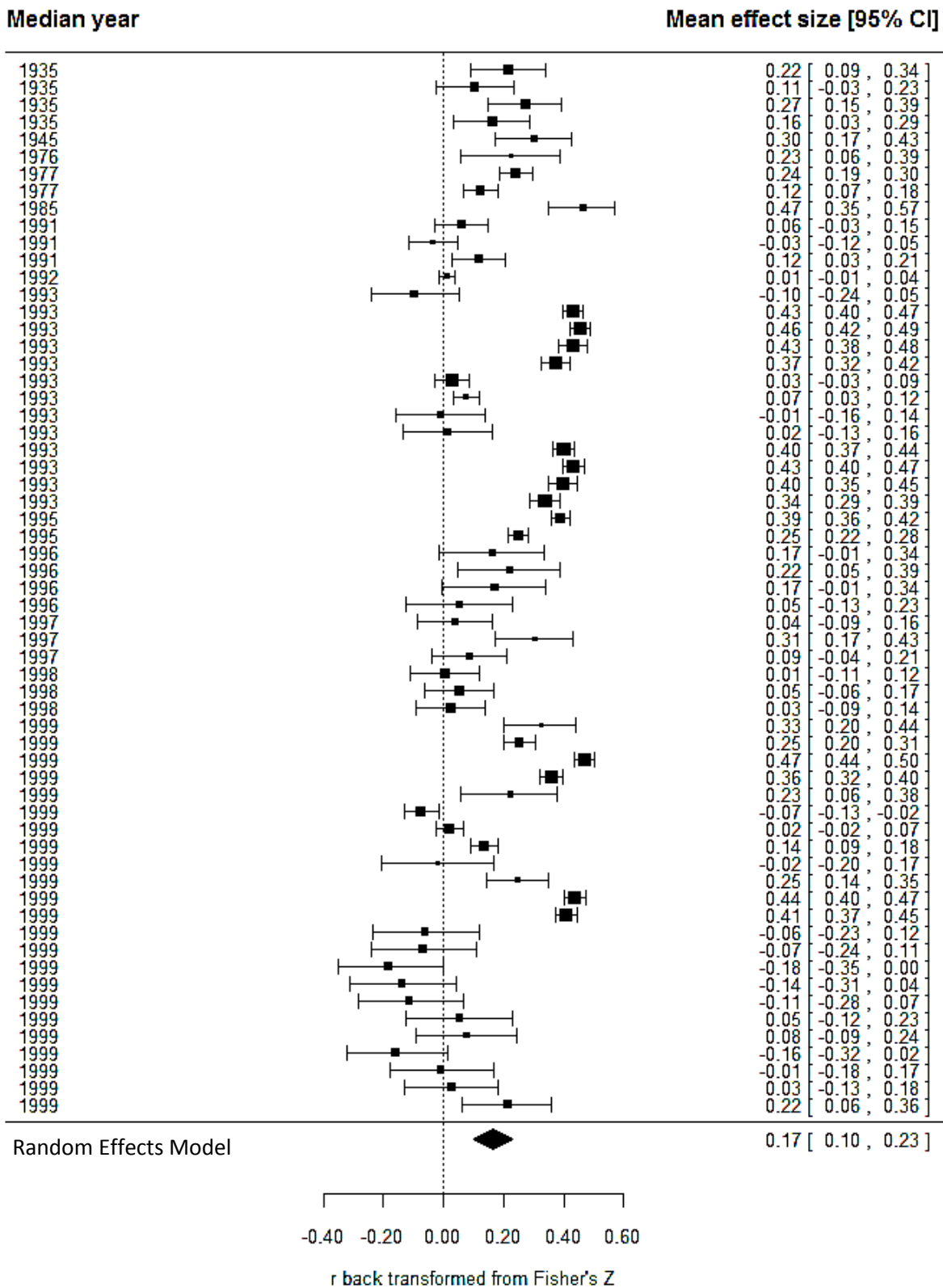
Fig. 3b Forest plot of stars' market brand equity- movie theatrical performance effect sizes sorted from 1999 to 2013
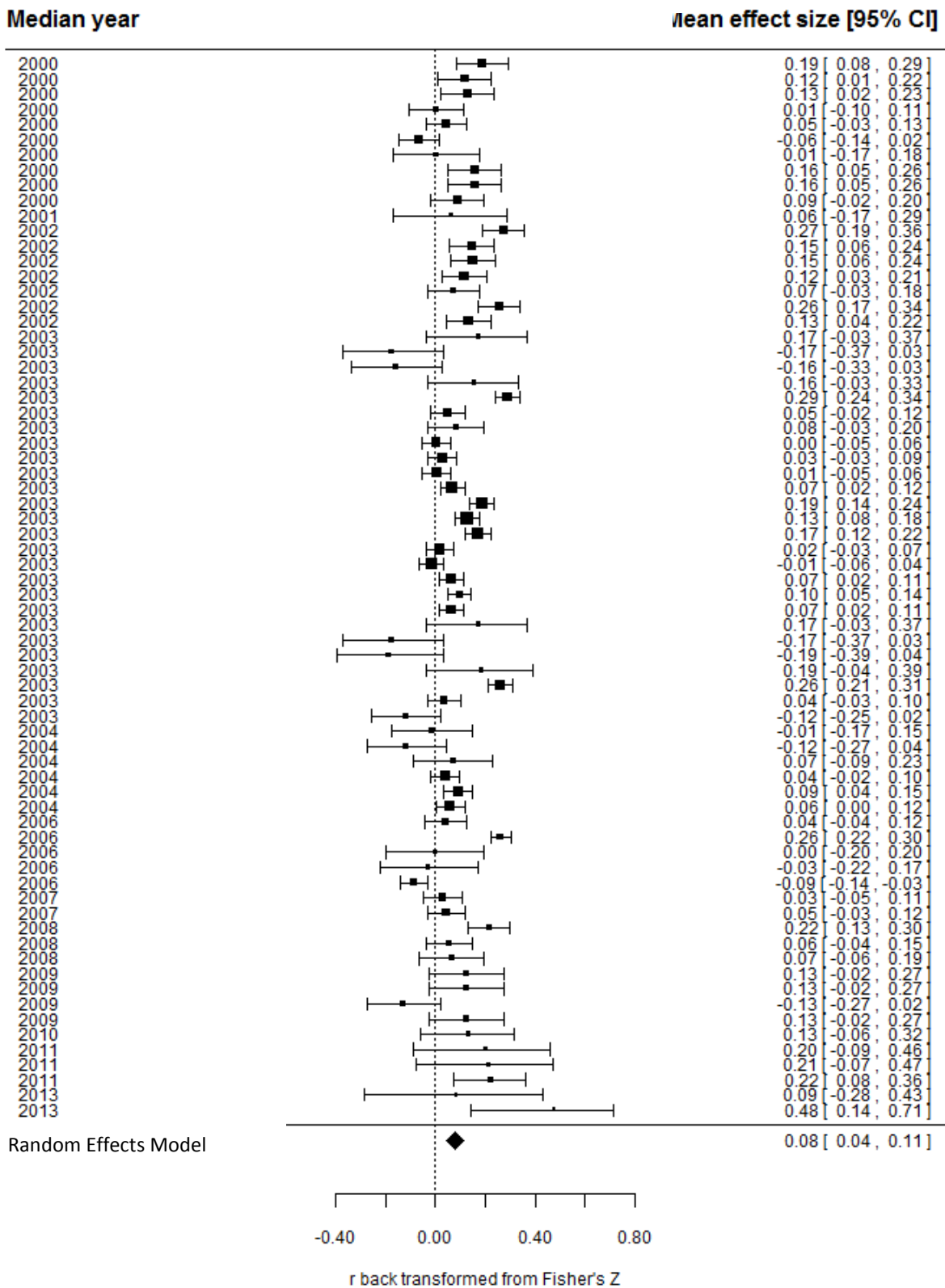
Fig. 4 Findings about star brand equity debate and assumption

Debate	Findings
1. Popularity versus artistic recognition. Does star popularity (market and media dimensions) or artistic recognition (artistic dimension) have a stronger impact on BO performance?	Star popularity, as per the market brand equity dimension, has a stronger impact on short-term BO than star artistic recognition (the artistic dimension). The media dimension of popularity has the same impact on short-term BO than star artistic recognition. Star popularity (market and media) and artistic recognition have the same impact on long-term BO.
Assumption	**Findings**
2. Fading star power. The impact of star brand equity on movie performance is weakening over time.	Fading star power is supported for star popularity (for both the market and media dimensions) but not for star artistic recognition (artistic dimension) which remains stable across the years.

Fig. 5 Findings about product reviews debate and assumptions

Debate	Findings
3. Influencer versus predictor effect. What is the impact of critics' reviews on short-term BO compared to long-term BO?	Critics exert a dual influencer and predictor role on BO revenue.

Assumptions	Findings
4. Users' advantage. Users' reviews are a stronger driver of BO performance than professional critics' reviews.	There is weak support for the users' advantage. The effect of the volume of users' reviews is stronger than the effect of the volume of professional critics' reviews on the full sample (the effect of their valence is of the same strength). However, their effects (either volume or valence) have the same strength when using only the primary studies in which both users' and critics' reviews were examined to strengthen statistical control.
5. Increasing user's advantage. The impact of users' reviews on BO performance is strengthening over time whereas the impact of critics' reviews is weakening.	There is no increasing users' advantage as the effects of users and professional critics' reviews are stable over time.