

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

# Electronic Commerce Research and Applications

journal homepage: [www.elsevier.com/locate/ecra](http://www.elsevier.com/locate/ecra)

## Towards effective online review systems in the Chinese context: A cross-cultural empirical study

Hui Fang<sup>a</sup>, Jie Zhang<sup>a</sup>, Yang Bao<sup>b,\*</sup>, Qinghua Zhu<sup>c</sup><sup>a</sup> School of Computer Engineering, Nanyang Technological University, Singapore<sup>b</sup> School of Computing, National University of Singapore, Singapore<sup>c</sup> School of Information Management, Nanjing University, China

### ARTICLE INFO

#### Article history:

Received 12 September 2012

Received in revised form 23 February 2013

Accepted 1 March 2013

Available online xxxx

#### Keywords:

Online review system

Electronic commerce

National culture difference

Cross-cultural study

Empirical analysis

### ABSTRACT

Online reviews, as one kind of quality indicator of products or service, are becoming increasingly important in influencing purchase decisions of prospective consumers on electronic commerce websites. With the fast growth of the Chinese e-commerce industry, it is thus indispensable to design effective online review systems for e-commerce websites in the Chinese context, by taking into account cultural factors. In this paper, we conduct two empirical studies on online reviews. Firstly, we study how culture differences across countries (i.e., China and the USA) impact the way in which consumers provide online reviews. Secondly, we investigate the impact of online reviews on product sales in the Chinese context, and show that directly copying the ideas of successful online review systems in the USA will deteriorate the effectiveness of the systems in China. Finally, we propose several suggestions for the development of effective online review systems in the Chinese context based on the results of our two empirical studies and the findings in previous studies.

© 2013 Published by Elsevier B.V.

### 1. Introduction

Nowadays, more and more people in China choose to shop in online stores rather than in traditional ones. According to the report by [China Internet Network Information Center \(CNNIC\) \(2012\)](#), the number of online shoppers in China has reached 194 million by the end of 2011. The proportion of online shoppers in total Internet users increased from 35.2% to 37.8% in the same period with a steady growth every year. These numbers demonstrate that the Chinese e-commerce industry has shown great potential through its fast development.

With the advent of web 2.0, users have been particularly productive in providing opinions (i.e., online reviews) of virtually every product ranging from televisions, cars, hotels, restaurants, books and movies to mobile apps ([Dellarocas 2003](#)). These online reviews, compared with those provided by merchandisers themselves, convey more consumer-oriented product information about purchase and usage experiences ([Lee et al. 2008](#)). On the other hand, it has long been recognized that these reviews are of great use to prospective users of these products ([Lee et al. 2008](#)), leading to substantial efforts being devoted to harnessing them. [Resnick and Zeckhauser \(2002\)](#) point out that consumers' online reviews have a direct impact on transactions. [Chevalier and Mayzlin \(2006\)](#) demonstrate that online

reviews have a significant influence on book sales on Amazon US ([www.amazon.com](http://www.amazon.com), the largest e-commerce website in the United States). Online merchandisers also have begun to embed consumers' online reviews in product advertisements in the form of complement to product introduction ([Lee et al. 2011](#)).

Due to the significance of online reviews, it is critical to design an effective online review system, where all the activities (i.e., review contribution and review harness) happen. In this paper, we aim to shed some light on how to develop effective online review systems in the Chinese context. It has been recently recognized that the effectiveness of online review systems in a country is also influenced by its national culture ([Koh et al. 2010](#)). However, many e-commerce websites in China tend to directly copy ideas from those world's leading ones in another country (e.g., the United States) without considering the culture differences. It has been proven that the ignorance of culture differences will cause the failure of those systems ([Bricks 2006](#)). Thus, it is important for e-commerce sites in China to consider culture differences seriously when adopting others' experience. Hence, the first thing we need to do is to investigate how culture across countries (China and the United States specifically) could impact the online review systems, and confirm that directly copying ideas of online review systems from the US context will not work well in the Chinese context, through exploring the culture differences between the US and China in terms of the behavior of providing reviews. Based on the confirmation, we then need to make clear how exactly online reviews theoretically impact consumers' purchasing behaviors, especially in the Chinese context. Without knowing of

\* Corresponding author. Address: COM2-01-02, 15 Computing Drive, School of Computing, National University of Singapore, Singapore 117417. Tel: +65-83003266; fax: +65-67797365.

E-mail address: [baoyang@comp.nus.edu.sg](mailto:baoyang@comp.nus.edu.sg) (Y. Bao).

this correlation, it is impossible to establish an effective online review system that could positively influence consumers' products choosing process, strengthen trust among consumers and boost product sales. We intend to especially investigate the correlation between online reviews and product sales in order to build a more effective review system. In the end, our findings, together with the comparative analysis of our findings with those of other studies in the US context (differences are explained by the results in our cultural study), can further provide significant implications for the development of effective online review systems in the Chinese context.

We conduct two empirical study in this paper. Our first study aims to investigate the effects of national culture on consumers' review providing behavior by exploring the review systems on Amazon US and Amazon China ([www.amazon.cn](http://www.amazon.cn)) respectively. In particular, we investigate the differences across countries on various properties of online reviews, and seek to theoretically explain these differences in the framework of dimensions of national culture proposed by Hofstede et al. (Hofstede 1980, 2001; Hofstede et al. 2010). We find that: (1) the way in which consumers provide online reviews is influenced by culture differences; (2) compared with those consumers in US, consumers in China tend to be less engaged in online review systems, provide less negative reviews, and value negative reviews more. In the second study, we establish a multi-dimensional review-related research framework, and create a panel data econometric model to explore the correlation between online reviews and product sales from the perspectives of number of reviews, quality of reviews, reputation of reviewers and the emotional tendency of reviews based on dataset collected from the Dangdang Books ([book.dangdang.com](http://book.dangdang.com), one of the largest Chinese online bookstores) (Fang et al. 2011). Our results show that: (1) number of reviews, emotional tendency of reviews and spotlight reviews are all positively correlated with product sales, while reviewer rank and length of reviews have negative effect on product sales; (2) the review voting mechanism of Dangdang is not significantly correlated with product sales. Based on the results of the two studies, we further provide several important guiding suggestions towards the development of effective online review systems in the Chinese context. These systems should pay attention to emotional tendency of reviews, reviewer ranking mechanisms, spotlight reviews, review format and review voting mechanisms.

The rest of the paper is organized as follows. Section 2 provides an overview of related works on online consumer reviews as well as those on the influence of national culture on e-commerce and online review systems. In Section 3, we provide a high level introduction of our methodology. We then describe the details of our two empirical studies in Sections 4 and 5 respectively. After that, we provide several suggestions towards effective online review systems in the Chinese context in Section 6. Finally, we conclude our current work and propose future work in Section 7.

## 2. Related work

In this section, we focus on reviewing previous works on exploring the impact of online reviews on products sales, and the inconsistent results regarding the impact of online reviews in the Chinese context and the US context. We then investigate the national culture differences between China and the US, which is considered as the most probable factor of the inconsistent findings about the effect of online review systems in previous studies, as well as the related cross cultural empirical studies in the literature.

### 2.1. Online consumer reviews

Previous research on online reviews mostly concentrates on six aspects: connotation (Hennig-Thurau et al. 2004), motivation

(Picazo-Vela et al. 2010), content (Ye et al. 2007, Korfiatis et al. 2011), effect (Chevalier and Mayzlin 2006, Chen et al. 2007, Park and Kim 2008, Utz et al. 2011), management (Hennig-Thurau et al. 2004, Park and Kim 2008), and individual differences (Chen 2008, Resnick and Zeckhauser 2002, Doh and Hwang 2005, Hu et al. 2011, Ba and Pavlou 2002). These aspects almost cover the whole life cycle of online consumer reviews.

Although there are many studies on the *effect* aspect focusing on the exploration of the impact of online reviews on the sales of products or service, only a few of them are conducted in the Chinese context. One of them is the work done by Hao et al. (2008), who have made use of movie panel data and emotional tendency of online reviews. Their findings include: (1) online reviews have different impact on movie box office at the different stages of movies' life cycle, and the impact reaches its peak after 3 weeks; (2) emotional tendency of online reviews plays a more significant role on movie box office than quantity of reviews; (3) positive influence of five-star reviews has overtaken the negative influence of one-star reviews; (4) there is no significant influence of reviews having two-star to four-star. Another study is conducted by Zhang et al. (2010), who aim to address the question of whether consumer-written-reviews and professional editor-written-reviews have different influences on the online consumers' purchasing intention. Based on the real data collected from dianping.com (one of the leading consumer advice websites in China), they find that, in the Chinese context, consumers' ratings on quality of food, environments and services of restaurants and the number of reviews are positively correlated with the online popularity of restaurants, whereas professional reviews have a negative relationship with consumers' purchasing intention. Different from these studies, our work in the Chinese context considers a more comprehensive framework, including different perspectives of online reviews such as review's quality and reviewer's rank, for the investigation on the impact of online reviews on book sales. Besides, we also explore the consumers' reviews providing behavior by taking the national culture factor into consideration. Thus, we expect to achieve reasonable solutions on the question of how to build an effective online review system in the Chinese context.

Some studies conducted in non-Chinese context lead to the inconsistent conclusions with those in the Chinese context. For example, Duan et al. (2008) discuss the effect of online reviews on movie box office from the perspectives of pervasive effect and awareness effect. The result turns out to be that online reviews cannot influence sales directly. Meanwhile, they use various methods to explore the correlation between product sales and different aspects of reviews. Chevalier and Mayzlin (2006) declare that online reviews have positive effect on book sales of both Amazon and Barnesandnoble.com. High quality reviews can improve sales, and reviews of one star have more significant effect on product sales than reviews of five stars, which is opposite to the study of Hao et al. (2008). Based on the Transaction Cost Theory and Uncertainty Reduction Theory, Hu et al. (2008) use a combinative analysis to exploit the effect of online reviews on sales of Amazon videos and books. Ghose and Ipeirotis (2010) use an econometric approach, text mining and technology of model predictive control to discuss the correlation between product sales and characteristics of reviews including average helpfulness of historical reviews, information richness and readability. They find that reviews mixed with objective sentences and subjective sentences have less impact on product sales than reviews containing only objective sentences or subjective sentences. Besides, Dellarocas et al. (2007) try to figure out how to predict movie box office based on historical reviews.

Overall, the research on the effect of online reviews on product sales results in many inconsistent conclusions in different culture contexts. Besides, most of the studies are conducted in the US

context (e.g., Amazon US), while a few have been carried out in the Chinese context. Consequently, most research results have only limited realistic value to guide e-commerce market in China. This motivates our work of comprehensively studying the impact of online reviews on product sales in the Chinese context. It should also be noted that the inconsistent results could be partly explained by culture differences between China and the US. The culture difference across countries can possibly result in consumers' different behavior in providing and deciphering reviews, and then bring about the difference of effectiveness of online review systems in different contexts. Therefore, in the next subsection, we examine the national culture difference across countries in the literature.

## 2.2. National culture difference

The values that distinguished countries from each other can be grouped statistically into six clusters. These six clusters become the Hofstede dimensions of national culture (Hofstede 1980, 2001; Hofstede et al. 2010), which serves as the most influential culture theory in the social science area. These six dimensions are power distance, masculinity–femininity, individualism–collectivism, uncertainty avoidance (Hofstede 1980), long term orientation (LTO) (Hofstede 2001), and indulgence versus restraint recently added (Hofstede et al. 2010). *Power distance* represents the degree to which the less powerful members of a society can accept unequally distributed power. People in countries with higher power distance accept a hierarchical order, while people in countries with lower distance power strive to equalize the distribution of power. *Masculinity–femininity* mainly expresses that masculine cultures emphasize work and material accomplishment, whereas feminine cultures weight more on human relationship. *Individualism* and *collectivism* are defined as a preference for a loosely-knit social framework and for a tightly-knit social framework respectively. People with higher individualism pay more attention to themselves and their immediate families than people with higher collectivism. The *uncertainty avoidance* dimension refers to the extent to which the members of a society feel uncomfortable with uncertainty and ambiguity, and take actions to avoid them. Societies with a *short-term orientation* exhibit a relatively small propensity to save for the future, and a focus on achieving quick results, while societies with a *long-term orientation* express a strong propensity to save and invest, and be perseverance in achieving results. *Indulgence* stands for a society that allows relatively free gratification of basic and natural human drives related to enjoy life and have fun. *Restraint* stands for a society that suppresses gratification of needs and regulates it by means of strict social norms.

Hofstede dimensions of national culture have already received strong empirical support from a great deal of studies (Hofstede and Bond 1984, Sondergaard 1994, Nakata and Sivakumar 2001, Hofstede et al. 2010). However, there are also some criticisms towards Hofstede's dimensions. For example, McSweeney (2002) points out that doubts about the representative of IBM population (questionnaire study subjects) is the most common criticism of Hofstede's dimensions. Besides, Dorfman and Howell (1988) criticize that the scale of power distance only focuses on national level which can not measure individual differences. We agree that Hofstede's dimensions of national culture does have its shortcomings and weaknesses. However, compared to other models on national cultures, it is a relatively representative and homogeneous cultural model which is improved over decades. Therefore, we choose to use it in our research as the reference of national culture differences between the US and China. Our comparison is more emphasized on culture differences on national level instead of individual level.

"Culture only exists by comparison" (Hofstede et al. 2010), and the US is chosen for our comparative study with China mainly

because of two reasons: (1) the US culture differs greatly from culture in China regarding to three of the six culture dimensions (see Table 1<sup>1,2</sup>); (2) e-commerce (e.g., Amazon and eBay) in the US has mature development, and most previous online review studies are also conducted on the Amazon platform in the US context. Thus, a detailed cross-cultural study can allow e-commerce sites in China to better learn from advanced experience in the US.

A few cross-cultural studies have been conducted on the topic of e-commerce or online consumer reviews. Pavlou and Chai (2002) apply the theory of planned behavior (TPB) to capture online shopping intentions in China and the US to address the question of what drives e-commerce across cultures. They incorporate Hofstede's cultural dimensions in their study, and conduct a questionnaire survey on 58 subjects from China and 55 subjects from the US. The empirical results manifest that culture difference plays an influential role in e-commerce adoption. From the perspective of behavioral theory, Koh et al. (2010) explore the correlation between online consumers rating behavior and cultures. According to the real data collected from IMDB.com (the US) and douban.com (China), they point out that consumers' rating behavior is influenced by different cultures. In conclusion, little previous research has studied the influence of national culture on the effect of consumer online review systems especially in terms of the detailed aspects of online review systems. Previous research did not examine how exactly the national culture influences the way of consumers in providing online reviews. Besides, the results of the above mentioned two studies can partly support that cultural difference does affect online review systems adoption among consumers across cultures, which further motivates our work.

## 3. Overview of the methodology

In order to scientifically achieve our research objectives, we construct our research by integrating approaches of literature study, econometrics methods, empirical approach and methodology of comparison. Our research mainly consists of two empirical studies, based on the analysis of online reviews data collected from Amazon US, Amazon China and Dangdang. The research methodology is described in detail as follows.

In the first study, we aim to explore the differences across countries on various properties of online reviews, such as the number of reviews, the length of reviews, helpful and unhelpful votes of reviews, and emotional tendency of reviews. More importantly, we seek to theoretically explain these differences in the framework of Hofstede dimensions of national culture. We firstly propose a set of research hypotheses based on our objective. Then, we examine the real data collected from Amazon US and Amazon China. The dataset of Amazon China includes 8669 review information of 644 novel books published in September 2011 and 49,935 review information of 89 books appeared in the 2011 bestsellers, and the dataset of Amazon US consists of 16,269 reviews of 500 novel books published in September 2011 and 76,973 review information of 106 books ever appeared in the 2011 bestsellers. Finally, based on the statistics data of the online reviews in the Chinese context and the US context, we test our proposed hypotheses by using a comparative study, where the two-sample *t*-test is used to check whether the differences of consumers' reviewing behavior between China and the US are significant.

<sup>1</sup> <http://geert-hofstede.com/countries.html>, 03/12/2012.

<sup>2</sup> There is less data about the *indulgence-constraint* dimension. Indulgence scores are highest in Latin America, parts of Africa, the Anglo world and Nordic Europe; restraint is mostly found in East Asia, Eastern Europe and the Muslim world (Hofstede and Hofstede). Due to the lack of direct comparison of this dimension between China and the US in the previous studies, we choose not to mention it in our cross-cultural study.

**Table 1**  
Cultural index scores of China and United States.

Country	Power distance	Individualism	Masculinity	Uncertainty avoidance	Long-term orientation
China	80	20	66	30	118
United States	40	91	62	46	29

In the second study, we investigate the effect of online consumer reviews on product sales in the Chinese context based on the panel data collected from Dangdang. First, we propose our research hypotheses, and establish a relatively comprehensive framework through lots of previous research results mostly towards Amazon US. These results will also be used to compare our results on Dangdang in the Chinese context. Second, a panel data model is constructed on the basis of econometric methods and strictly grounded on its basic procedures. We define product sales as a dependent variable and review related variables together with other correlated variables as independent variables, and then use the Pooled Ordinary Least Squared (POLS), Hausman test and Heteroscedasticity test consequently to verify the effectiveness of our research dataset and model. Third, the empirical study is conducted to explore the correlation between online reviews and product sales on the data collected from Dangdang. The data is crawled directly from Dangdang and covers basic book information and total review information of 12,609 new books posted from October 19, 2009 to April 16, 2010 within 180 days. We explore the impact of online reviews on consumers. Then, with the assistance of Stata 10.0, we apply the econometric model on dataset to test the proposed hypotheses and analyze the regression results.

Finally, we analyze the results of the two empirical studies, and compare with the results of the empirical research on Amazon US. It should be mentioned that we choose not to conduct the same study (as our second study) on Amazon US in this paper, mainly because there have been already a lot of results on Amazon US, which can be compared directly with our results on Dangdang. Through this process, we derive several important guiding suggestions on the aspects including emotional tendency of reviews, reviewer ranking mechanisms, spotlight reviews, review format and review voting mechanisms, towards the development of effective online review systems in the Chinese context.

## 4. Experiment I: a cross cultural study

### 4.1. Hypotheses

As stated, there exists a lot of cultural difference between the United States and China. We focus on studying how exactly the dimensions of national cultures are related to various properties of online reviews including the number of reviews, emotional tendency of online reviews, the length (word counts) of reviews, helpful and unhelpful votes of reviews, and spotlight reviews (eye-attracting reviews). In this section, we firstly propose hypotheses based on the analysis of inter-relationship between consumers and their online reviews with respect to different cultural dimensions in Hofstede's framework (Hofstede et al. 2010).

As argued by McCrae and Terracciano (2005), people in individualistic societies (e.g., US) are more likely to value freedom of expression and expected to be more talkative, extrovert and agreeable. In contrast, people in collectivistic societies (e.g., China) are more likely to seek group consensus, and people in countries with higher power distance (e.g., China) are recognized as introvert and disagreeable. We thus hypothesize that consumers in China are more reluctant to express their feelings on what they have experienced. In contrast, consumers in the US value freedom of expression more strongly (Koh et al. 2010), and hence feel themselves to be more free to openly comment on what they have

experienced, or convey their agreement or disagreement on reviews provided by others. They are also more willing to provide much more comprehensive and detailed reviews to express their appreciation or dissatisfaction than those in more collective culture (e.g., China). Therefore, by taking these two cultural dimensions (individualism and power distance) into consideration, we propose our first three hypotheses.

- H1:** The consumers in the US tend to provide a larger number of online reviews than consumers in China.
- H2:** The online reviews in the US context should receive a larger number of (un)helpful votes than those in the Chinese context.
- H3:** The reviews in the US context will be longer than those in the Chinese context.<sup>3</sup>

It is widely hold that China and the US have significant cultural difference on the Masculinity dimension (Fernandez et al. 2008). As shown in Table 1, the US has a lower index score than China, indicating that consumers in the US are more caring for others than those in China. We thus propose our fourth hypothesis as follows:

- H4:** Every online review in the US is more likely to be voted by review readers than that in China.

Consumers' evaluation of online reviews is also influenced by culture difference in terms of power distance, which refers to the extent to which people accept the unequally distributed power. In countries with higher power distance (e.g., China), less powerful objects are willing to depend on more powerful objects (Hofstede 2001). Analogously, consumers with little knowledge (less powerful) of products tend to have more reliance on spotlight reviews<sup>4</sup> (more powerful) provided by those with more information. Thus, we propose the fifth hypothesis as follows:

- H5:** Spotlight reviews can receive more helpful votes in China than those in the US.

Consumers' criticism of online reviews might be impacted by the culture difference in terms of uncertainty avoidance. The dimension of uncertainty avoidance is tightly correlated with neuroticism (McCrae and Terracciano 2005). According to Hofstede et al. (2010), people in weak uncertainty avoidance societies are more likely to be tolerant and moderate, while people in strong uncertainty avoidance countries tend to be more conservative and extreme. As shown in Table 1, China has a lower score (30) compared with that (46) of the US in this dimension. Therefore, we expect that Chinese tend to be more moderate to express their emotional tendency toward product or service. Besides, Hu et al. (2006) find that when consumers report their reviews towards product or service, under-reporting bias may exist. This is because consumers who are greatly satisfied or dissatisfied are more likely to provide reviews. Chinese consumers will provide less extremely

<sup>3</sup> The length of reviews is measured by word counts.

<sup>4</sup> A spotlight review of a product, usually put on the top of all the other reviews, means that the review has insightful content. It may be selected by the editors of websites (e.g., Dangdang), or refers to the reviews having a lot of votes from other users.



negative reviews than American consumers since Chinese consumers are in a weaker uncertainty avoidance society (China). Therefore, according to the culture difference on the dimension of uncertainty avoidance, we propose our sixth and seventh hypothesis as follows:

- H6:** Consumers in China generally hold a more positive attitude (by providing reviews with more positive emotional tendency) toward products or service than consumers in the US.
- H7:** The under-reporting bias for extremely negative reviews (1-star reviews) is less significant in the Chinese context than that in the US context.

The long-term orientation dimension is closely related to Confucianism, which favors risk aversion (Wang et al. 2005). China is extremely high on this dimension as shown in Table 1. Thus, we believe that Chinese tend to hold conservative attitude on risk and prefer to avoid and minimize risk towards the scenarios where their interests may be affected. Therefore, in online review systems, Chinese consumers will pay more attention to negative reviews and weight these negative reviews more in their decision-making process. Hence, we propose our eighth hypothesis as follows:

- H8:** The extremely negative reviews (1-star reviews) receive more helpful votes in China than those in the US.

## 4.2. Empirical analysis and results

### 4.2.1. Data collection

We collect online review data about books from Amazon US and Amazon China. Amazon US is one of the earliest online bookstores in the US, while Amazon China is one of the earliest online bookstores in China. Amazon China was founded in 2000, purchased by Amazon Corporation in 2004, and then was standardized to be as Amazon US (Wikipedia). We choose these two sites mainly because they have similar online review systems, which can well eliminate the effect of different systems on online consumers' reviewing behavior. For the purpose of a comprehensive study, we collect all the review information about two kinds of books: one type is new books under the category of novels published in September 2011 in the two sites,<sup>5</sup> and another type is the books ever appeared in the 2011 bestseller lists of these two sites. These two kinds are different, but can complement each other. The new books are in a relatively earlier stage of the sell life cycle, while others are in the mature stage. Therefore, the results derived with respect to these two different types of data can be more reliable and convincing. Although Amazon China is launched roughly 5 years later than Amazon US, we find that these two sites are comparable in terms of  $\frac{\text{the number of monthly unique visitors}}{\text{the number of book types}}$ , which is approximately 7.18 for Amazon US and 11.5 for Amazon China.<sup>6</sup> That is to say, on average, each book in Amazon US is expected to be browsed by seven unique people in a month, while 11 people in Amazon China. We can further infer that if people in two countries have the same tendency to give online reviews, each book in Amazon China should receive slightly larger volume of reviews than that in Amazon US.

Table 2 shows the statistic information of these two datasets. Note that Amazon discloses the information whether a reviewer has actually purchased the specific product. According to Table 2, we can see that a large part of reviews on Amazon US are provided

by users who did not purchase the product that they have commented. However, they may read or purchase the same book from other channels. For example, a reviewer may have bought a book in hardcover version, but he may also provide his opinion towards the same book in paperback version. On the contrary, almost all the reviews on Amazon China are provided by the users who actually have bought the corresponding books from the Amazon China site.

For each book, we crawl its basic information, including book ID, author(s), fixed price, sales price, publication date, publisher information and its sales rank till March 4, 2011. Besides, we also collect concrete review data of each book till March 4, 2011, including posted date, the number of stars of the review, reviewers' ID, reviewer's badges, review length (word counts), whether it is a spotlight review, whether it is a verified purchase, review text, helpful votes and unhelpful votes. Finally, we process the data to our needed format.

### 4.2.2. Data analysis and results explanation

Tables 3 and 4 summarize the results of our first comparative study. We apply the two-sample *t*-test to examine whether the difference between the review data of Amazon China and that of Amazon US is significant. We discuss our results based on the proposed hypotheses as follows.

We use the *average number of reviews per book*, *average number of votes per review* and *average word counts per review*<sup>7</sup> as the indicators for H1, H2 and H3 respectively. We can see that these three hypotheses are supported. In other words, consumers in the US is more talkative and extrovert, and are more engaged in the online review system compared to consumers in China. This is in accordance with the culture difference on the dimension of Individualism versus Collectivism between the US and China.

The *percentage of reviews having votes* is adopted to test our fourth hypothesis. Tables 3 and 4 indicate that reviews in China have less chance to be voted compared with those in the US. That is to say, consumers in China are less willing to evaluate other consumers' reviews, especially the unpopular reviews. Based on the results shown in the tables, we can see that the gap between the US and China in terms of this perspective is quite evident, but the difference of the masculinity index score between the US and China in Table 1 is not as significant as the indicated gap. We can infer that this may be caused by other cultural dimensions such as *individualism–collectivism* and *power distance*.

To test H5, we use the quantity of *relative average helpful votes per spotlight review*. For the new books, H5 is supported. Consumers in China pay more attention to spotlight reviews (the most eye-attracting reviews) than people in the US. However, for the best-seller books (see Table 4), H5 is not supported. This might be due to the effect of individualism–collectivism culture dimension. Specifically, consumers in a more collective society (e.g. China) tend to trust the (collective) selection of bestseller books more than those in a less collective society (e.g. US). Thus, they tend to have less uncertainty about the quality of the book as well as the reliability of its reviews. Consequently, the difference between spotlight and ordinary reviews are not that salient for them. In contrast, those who come from a less collective society tend to have uncertainty even if the book is selected as bestseller. They thus need more reliable reviews to help them to judge the quality of the book.

To test H6 and H7, we use *average stars per review* to measure online consumers' emotional tendency and *percentage of 1-star reviews* to measure under-reporting bias for extremely negative reviews. Both of them are supported (see Tables 3 and 4), indicating that consumers' review provision is influenced by the

<sup>5</sup> To avoid the multi-version problem on Amazon US, we limit our new books to be in paperback version.

<sup>6</sup> Nearly in the same period of 2011, the number of monthly unique visitors for Amazon US is around 282,000,000 (comScore), while around 29,810,000 (iResearch) for Amazon China. We search on the Amazon US and Amazon China respectively in the same day (July 18, 2012), and find that there are around 39,273,988 types of books sold on Amazon US, while 2,599,128 sold on Amazon China.

<sup>7</sup> It should be noted that we use the word counts instead of bytes of reviews to represent the length of reviews. It is said that for the same content, the number of Chinese words is equivalent to 1.6 times of the number of English words. (Language)

**Table 2**  
Statistical information about the two datasets.

Dimension	New books		Bestseller books	
	United States	China	United States	China
Number of books	500	644	106	89
Number of reviews	16,269	8669	76,973	49,935
Percent of reviews from purchased consumers	38.35%	84.76%	50.91%	91.77%
Average time duration since publication (months)	6	6	17.8	23.2

**Table 3**  
Comparison of aspects of online reviews on Amazon new books between China and the US.

Review aspects	United States	China	Difference degree <sup>a</sup>	$Pr( T  >  t )$	Whether significant	Testing hypothesis	Outcome
Average Number of Reviews per Book	32.5380	13.4619	1.4170	0.0000	Yes	<b>H1</b>	Passed
Average Number of Votes per Review	5.5546	0.5309	9.4626	0.0000	Yes	<b>H2</b>	Passed
Average Word Counts per Review	168.5368	65.8445	1.5596	0.0000	Yes	<b>H3</b>	Passed
Percentage of Reviews Having Votes	70.66%	16.96%	3.1663	–	Yes	<b>H4</b>	Passed
The Relative Average Helpful Votes per Spotlight Review	8.6157	11.7277	–0.2654	0.0000	Yes	<b>H5</b>	Passed
Average Stars per Review	3.9569	4.3318	–0.0866	0.0000	Yes	<b>H6</b>	Passed
Percentage of 1-star Reviews	8.86%	3.31%	1.6767	–	Yes	<b>H7</b>	Passed
The Relative Average Number of Helpful Votes for per 1-star Review	1.9166	2.1973	–0.1278	0.0000	Yes	<b>H8</b>	Passed

<sup>a</sup> It is always computed by  $(US - China)/China$ .

**Table 4**  
Comparison of aspects of online reviews on Amazon bestseller books between China and the US.

Review aspects	United States	China	Difference degree	$Pr( T  >  t )$	Whether significant	Testing hypothesis	Outcome
Average number of reviews per book	726.1604	554.8333	0.3088	0.0000	Yes	<b>H1</b>	Passed
Average number of votes per review	8.9586	0.6051	3.8052	0.0000	Yes	<b>H2</b>	Passed
Average word counts per review	138.6407	48.6865	1.8476	0.0000	Yes	<b>H3</b>	Passed
Percentage of reviews having votes	72.81%	11.42%	5.3757	–	Yes	<b>H4</b>	Passed
The Relative average helpful votes per spotlight review	107.702	71.1631	0.5135	0.0527	Yes	<b>H5</b>	Failed
Average stars per review	4.1376	4.4091	–0.0616	0.0000	Yes	<b>H6</b>	Passed
Percentage of 1-star reviews	9.41%	2.27%	3.1454	–	Yes	<b>H7</b>	Passed
The relative average number of helpful votes for per 1-star review	2.7748	3.2984	–0.1587	0.0000	Yes	<b>H8</b>	Passed

uncertainty avoidance cultural dimension. In other words, the reviews provided by consumers in China are more positive, and the under-reporting bias for negative reviews is less significant compared with online reviews provided by consumers in the US.

We examine the *helpful votes per extremely negative reviews* to check consumers' attitude towards risk in online review systems. As shown in Tables 3 and 4, consumers in China are more risk-aversion. They value negative reviews more heavily compared to users in the US. That is to say, consumers' reviewing behavior is influenced by the long-term orientation dimension.

To summarize, consumers' online reviewing behavior is influenced by culture differences with regard to five of the six dimensions of national cultures, but the effect of power distance with regard to spotlight reviews can become smaller as books become more popular online. Compared to consumers in the US, consumers in China are less engaged in the online review systems and "helpfulness" voting mechanism. Besides, Chinese tend to provide positive reviews towards books, provide less extremely negative reviews, and pay more attention to the negative reviews provided by other online consumers.

## 5. Experiment II: panel data econometric model

In this section, we conduct our second empirical study. We first propose hypotheses with respect to the correlations between the

aspects of online reviews and product sales. We then test these hypotheses by fitting a linear panel model to our collected dataset.

### 5.1. Hypotheses

Online sales or consumers purchasing decision making might be affected by many factors, which have been explored by researchers in psychology, behavioral science, economics and other fields. From the demographic, cognitive and psychological aspects, Wang (2001) tries to figure out which sort of potential consumers would actually purchase online. They find that online consumers' attitude and intention are greatly influenced by cognitive and psychological characteristics, but not affected by demographic features. Therefore, we assume the following premise for our study: products' demographic characteristics have little effect on products' sales. We then study the correlation between online reviews and product sales from the dimensions elaborated in the following.

Online consumers can know the number of reviews about one particular product. As discussed in Section 2, a lot of empirical analysis have proven the positive impact of the number of reviews on product sales. Generally speaking, the larger number of reviews for a product, the more information about the same product can be acquired by potential consumers, and this can be more beneficial for consumers to know about the product. Godes and Mayzlin (2004) point out that the more discussions on a product, the more

likely for it to receive intensive attention which could lead to more purchase of this product. Hence, we propose our first hypothesis:

**H1:** The number of reviews has a positive impact on product sales for the same kind of products.

A lot of online review systems provide reviews of different ratings. For instance, reviews are rated one to five stars on Dangdang. The different ratings represent feeling of likes or dislikes toward the particular product or service, and they also reflect different levels of persuasive effect. No matter positive reviews or negative reviews, they are all from the users' angle to describe their experience of using the product, and the reviews can include the content and quality characteristics of the products. Then, based on provided information, the review receivers evaluate the product, compare with their expectations and then decide whether to purchase the product. Following the research of Chevalier and Mayzlin (2006), we hypothesize that positive reviews can have positive effect towards potential consumers, while negative reviews can have negative effect. Therefore, we propose our second hypothesis:

**H2:** The emotional tendency of online reviews has significant effect on product sales. Generally speaking, the higher overall rank of product reviews, the more sales the product would receive.

Reputation is what is generally said or believed about a person's or thing's character or standing (Josang et al. 2007). In an online review community, connection between information senders and receivers is very weak, which is completely different from traditional face-to-face communication. In order to eliminate negative factors caused by the new form of online communication in product transactions, e-commerce websites designers attempt to introduce the concept of the reviewers' rank, which is determined by several pre-set factors. Consumers can decide whom can be trusted or distrusted by viewing the rank of every reviewer. On the basis of the number of review medals, Dangdang divides its reviewers into five classes: Primary, Intermediate, Advanced, Premium and Expert.<sup>8</sup> Review medals can be obtained either by posting a review more than 50 words or when posted reviews have been chosen as spotlight reviews on the website. The former research has elaborated that reviews posted by reviewers of high reputation can greatly affect book sales. Guadagno and Cialdini (2003) point out that authority is far more important in the process of online decision-making than in traditional face-to-face interaction, and online product reviews are in the form of indirect communication. We expect that the higher rank (reputation) of a reviewer, the much more important effect of her review on product sales. Therefore, we propose our third hypothesis as follows:

**H3:** The reputation of a reviewer is significantly and positively correlated with product sales.

Quality of a review refers to its authenticity, reliability, relevance of its content with the product and whether it provides helpful information for potential consumers. We test the effect of review's quality on product sales from the aspects of review votes, length and whether it is a spotlight review.

In Dangdang's review system, people can comment on existing reviews and vote whether the reviews are helpful. Therefore, we can use votes of a review as indicator to the quality of the review. Chen et al. (2007) point out that for votes in a range of either

"helpless" or "helpful", reviews with the votes of "helpful" can promote product sales. In contrast, reviews with the votes of "helpless" may hinder product sales. Therefore, we propose our fourth hypothesis:

**H4:** The votes of "helpful" for reviews are significantly and positively correlated with product sales.

Review length can also be used to measure the quality of a review. Intuitively, a longer review tends to contain more information. Consequently, consumers are more likely to obtain their needed information, making themselves be confident to make a purchasing decision. Hence, we propose our fifth hypothesis:

**H5:** Review length is positively correlated with product sales.

Consumers' information search cost is a key factor that influences online product sales. The spotlight review mechanism of Dangdang is similar to that of Amazon. The spotlight reviews are selected by the editors of Dangdang. When they think that a review has unique perspectives, insightful content, personalized text, and more importantly a lot of "helpful" votes, they would mark it as a spotlight review during a particular period. The high availability of spotlight reviews means low information search cost, so they would have much effect on potential decision-making of consumers. A lot of research towards e-markets demonstrates that majority of consumers think that the information processing cost is relatively high, and ordered information has a disproportionately strong influence on consumer decision-making (Smith and Brynjolfsson 2001). Therefore, we propose our sixth hypothesis:

**H6:** Spotlight reviews have more positive impact on product sales than other reviews.

Based on the above mentioned research hypotheses, we build our research framework as shown in Fig. 1.

## 5.2. Empirical analysis and results

In this section, we systematically analyze real data of online reviews and product sales, and present our study results.

### 5.2.1. Research object

The data of our research is collected from Dangdang Books and we choose Dangdang based on the following three reasons. First, Dangdang is the largest online Chinese bookstore in the world. Dangdang, as a typical and successful Chinese B2C website, is usually compared with Amazon both in academic and in business circles (Tang 2008). Our research tries to compare our study results towards Dangdang with former studies based on Amazon from the perspective of online book reviews to explore our results more comprehensively in the Chinese context. We then use these results to guide online review system building of Chinese e-commerce websites. Second, Dangdang has a comprehensive book review system. On

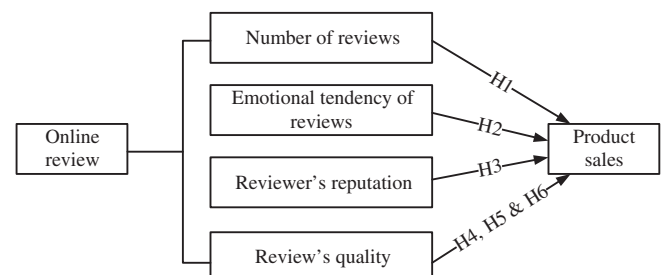


Fig. 1. Research framework of the effect of online reviews on product sales.

<sup>8</sup> In our research model of Section 5.2.3, these five classes are mapped into the numbers of "1", "2", "3", "4" and "5", respectively. A larger number refers to a higher rank of a reviewer.

Dangdang, consumers may not only read the reviews posted by others, but also comment on these existing reviews and vote for their quality. Furthermore, reviewers are also ranked. All these aspects are necessary for our research. Third, we can obtain the exact book sales data on Dangdang. Through our prior study of Dangdang, we find that it will conserve its daily sales data over a half year period and all these sales data can be collected by our own designed crawler. This is a special part of our research compared with former studies on Amazon. Because Amazon can only conserve its book sales ranking data, researchers (Chevalier and Mayzlin 2006, Chen et al. 2007, Duan et al. 2008, Brynjolfsson et al. 2003) basically treat sales ranking as an alternative variable of sales when exploring correlation between product sales and online reviews, which is following the research of Schnapp and Allwine (2001). The results they published in UCB/SIMS 2001 indicate that relation between book sales and sales ranking on Amazon approximately conforms to Pareto distribution, meaning that the logarithm of book sales is linear correlation with logarithm of sales ranking. However, since Pareto distribution is just approximate and that research was based on the data collected before 2001, book sales and sales ranking could not be simply treated as the same thing and using this relationship might influence the accuracy of research results.

### 5.2.2. Data collection

We collect our data from Dangdang covering all the new books posted from October 19, 2009 to April 16, 2010 within 180 days. In all, the dataset has 12,609 books, which includes basic information, daily sales and reviews related data of books. There are two reasons for studying new books. On one hand, consumers have not formed complete views towards new books yet, which means that online reviews can have a much greater effect on influencing book purchase. On the other hand, the increase rate of the reviews for new books is relatively higher.

With regard to each book, we use our own crawler to retrieve their basic information, including book category, identification, fixed price, sales price, publication data, publisher information and its daily sales since posted online. Besides, we also collect concrete review data of each book, including posted date, the number of stars of the review, reviewer's ID, reviewer's rank, review length and whether it is a spotlight review.

In the end, we sort, clear, integrate and standardize our data and pre-process it assisted by Stata 10.0, to make sure that it is in the form needed by our research.

### 5.2.3. Research model

For each book  $i$  sold on Dangdang, we take its book sales as the dependent variable. We use the logarithm form of the variable book sales since book sales conform to scale effect, and logarithm of book sales can well reflect the effect of independent variables on dependent variables. Scale effect is equivalent to that the majority of consumers might view popular books on sales while just a few consumers might view unpopular books. Therefore, the number of consumers conducting actual purchase actions can be treated as the function related to online reviews. By following the research of Chevalier and Mayzlin (2006) and Chen et al. (2007) and considering the real situation of Dangdang, we establish our econometric model shown as follows:

$$\ln S_i^t = v_i + \alpha \ln P_i^t + \beta_1 \ln R_{ni}^t + \beta_2 R_{si}^t + \beta_3 R_{ri}^t + \beta_4 R_{li}^t + \varepsilon_i^t$$

In the model,  $S_i^t$  is the sales of book  $i$  at time  $t$ ;  $P_i^t$  is the price of book  $i$  at time  $t$ ;  $v_i$  is the individual effect of book  $i$ .  $v_i$  does not change with time and is generally related to factors which cannot be directly observed or easily quantified;  $R$  is a vector related to online reviews, where  $R_{ni}^t$ ,  $R_{si}^t$ ,  $R_{ri}^t$ ,  $R_{li}^t$  respectively represent number of reviews, average review stars, average reviewer's rank and aver-

age length of reviews towards book  $i$ , and  $\beta_1, \beta_2, \beta_3, \beta_4$  are just mapping to their correlation coefficient;  $\varepsilon_i^t$  is the residual;  $t$  refers to a specific time (i.e. time point after aggregation). Following the previous work of Duan et al. (2008), we do not consider the lagged effect of online reviews at time  $t - 1$  on sales. This is mainly because in our research, each time point involves a relatively long period (15 days), and online reviews on Dangdang are in a default order of time descending. In this case, reviews of the current time point are much more influential than those of the previous time point on current sales. At the data processing stage, all the variables have been standardized to be in the same order of magnitude. It should be noted that although there are four shipping methods available on Dangdang, consumers may not be significantly different on shipping methods, which is caused by shopping habits of Chinese people. Hence, in this model, we would not consider the effect of shipping time on book sales. We may adjust our model in our future research if needed according to actual situations.

Our panel data consists of the observations of 12,609 units (books) across 12 time points. The panel data model is chosen based on the following three reasons: (1) the dummy variables can be set to control individual differences; (2) by combining the different values at different observation points of different cross-sectional units, we can increase the degree of freedom (DOF) and reduce the effect of collinearity among explanatory variables to improve the effectiveness of estimated results; (3) by repeatedly observing the same unit, we can better analyze dynamic changes of economic behaviors.

Then, we conduct Pooled Ordinary Least Squared (OLS), Hausman test and Heteroscedasticity test consequently to verify the effectiveness of our research dataset and model. Firstly, the OLS approach is used to verify the quality of our panel dataset. By assuming panel data as normal cross-sectional data and then performing regression analysis, and assisted by Stata 10.0, we find that the  $R$ -square value is 0.9023. It means that the dependent variable can be over 90% explained by independent variables, indicating that our panel dataset is of high quality and our model well coordinates the overall relationship between the dependent variable and the independent variables.

Secondly, Hausman test is applied to determine which form of panel model is more appropriate for our research: fixed effect model or random effect model. The Hausman test value is 8212.69, and by referring to critical chi-square distribution table, we find that the Null hypothesis is not supported, so the fixed effect model is more suitable for our research. The result can be explained as follows. Generally speaking, a particular kind of books is produced in high volume before actually entering the market and their sales are mainly determined by their downstream retailers, but their marginal price set by retailers is little related to their sales. Therefore, we can assume that their supply is fixed and not influenced by other external factors such as price set by retailers. Moreover, inherent difference exists in each book and is not time-dependent because of the difference among readers and topics.

Conducting Heteroscedasticity test is the last step of model testing and we find that our model has the cross-sectional heteroscedasticity. Related to our research object (i.e. books and reviews), heteroscedasticity can be explained from two aspects. On the one hand, books vary about their categories and are subjective products, which leads to great disparity of their topics and readers. Therefore, differences among observation values of dependent variables are deemed to increase deeply. On the other hand, heteroscedasticity is known to mostly exist in cross-sectional units. As we mentioned before, our dataset includes 12,609 samples but only 12 time observation points, which significantly increases heteroscedasticity of the model. In our research, the Heteroscedasticity Robust and BOOTSTRAP approaches will be used to adjust standard error, heteroscedasticity and serial correlation.



**Table 5**  
Regression results of the fixed effect model.

$S_i^t$	Coefficient	Standard error	$t$	$P >  t $	[95% conf. interval]	
$P_i^t$	<b>0.9857472</b>	0.0010385	949.19	0.000	.9837117	.9877827
$R_{ni}^t$	<b>0.4273748</b>	0.0045244	94.46	0.000	.4185070	.4362426
$R_{ri}^t$	<b>-0.1302111</b>	0.0061536	-21.16	0.000	.1422720	.1181502
$R_{si}^t$	<b>0.0396388</b>	0.0014393	27.54	0.000	.0368178	.0424598
$R_{li}^t$	<b>-0.0312003</b>	0.0007497	-41.624	0.000	-.0326697	-.0297308
$v_i$	0.0494393	0.0016667	29.66	0.000	.0461726	.0527060
sigma_u	0.33396631					
sigma_e	0.49688750					
rho	0.31117174 (fraction of variance due to $u_i$ )					

**Table 6**  
Regression results of the extended model.

$S_i^t$	Coefficient	Standard error	$z$	$P >  z $	[95% conf. interval]	
$P_i^t$	<b>-0.0877339</b>	0.0179895	-4.88	0.000	-.1229927	-.05247510
$R_{ni}^t$	<b>0.3957532</b>	0.0091027	43.48	0.000	.3779123	.41359410
$R_{ui}^t$	<b>0.0000673</b>	0.0021012	0.03	0.974	-.0040510	.00418550
$R_{wi}^t$	<b>0.0092465</b>	0.0023506	3.93	0.000	.3779123	.41359410
$R_{si}^t$	<b>0.0281511</b>	0.0024203	11.63	0.000	.0234073	.03289935
$R_{li}^t$	<b>-0.0067945</b>	0.0047524	-1.43	0.153	-.0161089	.00252000
$R_{ri}^t$	<b>-0.0103086</b>	0.0098482	-1.05	0.295	-.0296107	.00899350
$v_i$	1.874451	0.0593817	31.57	0.000	1.758065	1.9908370
sigma_u	0.46467211					
sigma_e	0.64914187					
rho	0.33880176 (fraction of variance due to $u_i$ )					

#### 5.2.4. Data analysis and results explanation

According to our proposed fixed effect model, we conduct the regression analysis of our sample data by using Stata 10.0 and the results are shown in Table 5. By viewing the corresponding  $P$  values, we find that all independent variables are significantly correlated with dependent variable. One point that should be emphasized is the positive effect of the price on book sales. Among all the factors that affecting book sales, price is in a very prominent position, which can also be seen in Table 6. This is inconsistent with Amazon's result where the coefficient between book sales and price is from  $-0.5249$  to  $-0.8132$  (Chevalier and Mayzlin 2006). The differentiated results might be caused by the fact that only new books are sold on Dangdang and lowest price strategy is used to cumulate loyal consumers, while on Amazon both new version and old version books are sold, which may increase the individual differences among books.

Meanwhile, reviews' stars and number of reviews are positively correlated with book sales, while length of reviews and reviewer rank are negatively correlated with book sales. We discuss our results based on the proposed hypotheses as follows:

##### 1. Number of reviews vs. book sales

We can see that the first hypothesis is supported (see  $R_{ni}^t$  in Tables 5 and 6), i.e. the number of reviews has a significantly

positive effect on book sales. Books are subjective products and as the number of reviews increases, it will be more possible for consumers to obtain helpful information on the products they want to know, and then promote more sales. This can be explained by Matthew Effect and Population Effect in Behavior Science (Merton 1968) where relatively more attention will lead to more and more attention, and then initiate Eyeball Effect.

##### 2. Emotional tendency of reviews vs. book sales

As shown in Table 5, review stars  $R_{si}^t$  are positively correlated with product sales and the correlation coefficient is 0.04. Our second hypothesis is verified and the emotional tendency of online reviews has significant effect on product sales. Generally speaking, the higher the overall rank of product reviews is, the more product sales would be. The reviews differing from one-star to five-star on Dangdang all reflect emotional tendency of readers towards particular books and in describing the content and plot of books. Moreover, they may comment the books from several concrete aspects of books such as styles and novelty. The potential consumers can decide whether books have met their expectation and whether to buy by viewing reviews posted by others. Oppositely, the second hypothesis is not supported in majority of former studies (Chen et al. 2004, Duan et al. 2008). The result indicates that emotional tendency of

reviews, either positive or negative, is highly valued by Chinese consumers especially for subjective products and can greatly affect purchase decision making process of potential Chinese consumers.

### 3. Reputation of reviewers vs. book sales

By considering both registration time and active degree of online participation of reviewers, Dangdang divides its reviewers into five classes, which are from primary reviewers to expert reviewers. Our empirical study shows that reviewer rank is negatively correlated with book sales (see  $R_{ri}^t$  in Tables 5 and 6). Thus, the third hypothesis is not supported, which can be partly explained by the following reasons. In real world, people are more willing to trust the authorities in a particular domain. For example, people would highly trust gourmet about good restaurants and food. However, a reviewer with relatively high rank does not mean that he is an expert in the online community of Dangdang. The longer online time and the more active participation give higher rank to a reviewer, but this cannot assure the quality of the reviews posted by the reviewer. More importantly, this kind of design may only reduce the degree of persuasive effect of the reviewer's reviews on purchase decision making of potential consumers.

### 4. Quality of reviews vs. book sales

To have more comprehensive exploration between quality of reviews and book sales, we add two independent variables, the number of "helpful" votes for reviews ( $R_{ui}^t$ ) and spotlight reviews ( $R_{wi}^t$ ) to our econometric model and the additional regression analysis results are shown in Table 6. We choose to use two models mainly for two reasons: (1) our POLS study shows that the four review-related variables in the first model can explain over 90% of our dependent variable, demonstrating that the model can well fit the observed data; and (2) different from the other four independent variables,  $R_{ui}^t$  and  $R_{wi}^t$  is infeasible to be aggregated into 12 time points (15 daily observation/time point). This is because there are only two spotlight reviews for each unit (i.e. book) and only a few reviews are associated with helpful vote. Therefore, we further examine these two variables in the second model which uses original 180 daily observations instead of 12 aggregated observations. The major differences of the two models are summarized as follows: (1) the first model in Table 5 is a fixed effect panel data model, while that in Table 6 is a random effect panel data model. To choose between fixed effect and random effect is determined by the Hausman test; and (2) the panel data used in the first model is balanced with 12 time points, while in the second model, the panel data is unbalanced, since there are some time points at which we have no observations of some units (books).

**Length of review:** According to  $R_{ri}^t$  in Tables 5 and 6, we find that our fifth hypothesis is not supported, and length of reviews is negatively correlated with book sales, although the absolute value of correlation coefficients is low. The results can be explained by the following three reasons. First, conceptually, potential consumers can gain more useful information from longer reviews. But long reviews also sort of scale up emotional tendency (either positive or negative) of potential consumers. This will lead to a relatively moderate result on effect of long reviews on product sales. Second, the longer a review is, the more time potential consumers will take to view it and gain useful information. This is inconsistent with the objective of low information search cost. Third, a long review might have more redundant information than a short one. Therefore, the value of it should be underestimated if website designers or managers have not regularized reviews.

**Spotlight reviews:** Spotlight review mechanism on Dangdang is similar to that on Amazon, and they are selected by content editors of Dangdang. As shown in Table 6, our sixth hypothesis is

supported that spotlight reviews is positively correlated with book sales. This meets our expectation mainly because spotlight reviews are more visible to potential consumers, take them less information search cost, and then have much greater impact on purchase decision-making. Potential consumers are more willing to view reviews which are more visible, location significant and time-saving.

**"Helpful" votes of reviews:** "Helpful" votes of reviews are not significantly correlated with book sales, which can be concluded from Table 6, so our fourth hypothesis is not supported. It means that the review voting mechanism does not show its value on Dangdang's online review system. Although this mechanism is also similar to that on Amazon, it is not actually active and useful due to the lack of effective guidance. Through browsing the online review community of Dangdang and viewing our sample data, we find that few people have voted for the quality of reviews, and commenting on existing reviews is much rarer.

## 6. Discussion

In this section, we propose several suggestions on the development of online review systems in the Chinese context based on our findings of the two empirical studies and literature review of others' studies as shown in Table 7. The detailed explanations will be elaborated in the following paragraphs.

### 6.1. Evaluating rank of reviewers

Although the empirical study by Chen et al. (2007) on Amazon US shows that reviewer rank is not significantly correlated with product sales, our second study (H3) demonstrates that reviewer rank is negatively correlated with product sales on Dangdang. These two results both demonstrate that reviewer rank does not reflect reviewers' expertise, because the extant ranking of reviewers is mainly based on reviewers' online time and the number of reviews posted. The content and quality of their reviews are rarely considered.<sup>9</sup>

We suggest that Chinese online review systems should pay more attention on improving their reviewer ranking mechanism. Specifically, the content of reviews should be taken into account when ranking reviewers. For instance, if there exists an indicator showing whether reviews of a particular reviewer are helpful or not, rank of reviewers might be more persuasive for potential consumers' decision making. This indicator's payoff will even be more salient in the Chinese context than that in the US context, due to the moderator effect of culture (i.e., power distance) – China is a society with higher power distance (Hofstede 2001) as indicated by H5 of our first study.

### 6.2. Guiding emotional tendency of reviews

Ye et al. (2007) point out that emotional tendency of online reviews will have a significant effect on movie box office. Book is similar with movie in the sense that they are both subjective products. Our second study (H2) demonstrates that emotional tendency of online reviews is positively correlated with product sales. That is to say, reviews on books, which reflect subjective feelings of readers, can also influence potential consumers' purchase intention in the form of seeking sympathy and recognition from others of the same interest. The review readers will decide whether a book meets their expectation and then make their final decisions. As

<sup>9</sup> It is worth mentioning that recently Amazon US has in certain degree improved its mechanism of this problem by considering the helpful votes that each reviewer has ever received, from which Dangdang should learn.

**Table 7**

Summary of suggestions based on Experiments I and II.

Perspectives of online reviews (Experiment II)	Culture implications (Experiment I)	Suggestions
Reviewer rank is negatively correlated with product sales on Dangdang (H3)	China is a society with higher power distance (H5). Online review from a reviewer of higher rank is more powerful	Improve the evaluation on rank of reviewers
Emotional tendency of online reviews is positively correlated with product sales (H2)	(1) Online consumers in China tend to provide positive reviews even when they are not satisfied and avoid to provide extremely negative reviews (H6). (2) Negative reviews can induce more consumers to engage in the voting system in the Chinese context (H7)	Encourage users to express their true feelings via online reviews
A longer review on Dangdang cannot promote more sales (H5)	(1) Online consumers in China are not self-motivated to provide online reviews (H1). (2) Chinese are likely to provide much shorter reviews than consumers in the US (H3)	Standardize format of reviews
"Helpful" votes of reviews are not significantly correlated with book sales (H4)	Only a limited proportion of online reviews has ever received votes from other users in China (H4)	Strengthen and promote review voting mechanism
Spotlight reviews have more positive effect on product sales than other reviews (H6)	Consumers in China are more willing to follow the spotlight reviews (H5), encouraging engagement in review systems	Further improve spotlight reviews
Number of reviews is positively correlated with product sales (H1)	Online consumers are not self-motivated to participant in online review systems (H1), and reluctant to provide reviews (H2) and votes for others' reviews (H4)	Encourage consumers rightharrow participate in online communities

mentioned before, this kind of behavior is particularly prevalent in Chinese online bookstores.

However, as studied in our cross-cultural study (*H6 and H7*), online consumers in China tend to provide positive reviews even when they are not satisfied and avoid to provide extremely negative reviews. Therefore, an effective review mechanism in the Chinese context is necessary to encourage readers to express their true feelings about books. This mechanism should encourage more potential consumers to be involved in providing reviews, and thus can expand consumer group both from the dimensions of depth and breadth. Most importantly, since negative reviews can induce more consumers to engage in the voting system in the Chinese context (*H7* of the first study), online consumers should be encouraged to truthfully express their negative attitude towards products or service. Negative reviews should be highly estimated and be processed legally, and they also should not be distorted in a reasonable range.

### 6.3. Quality of reviews

First, standardize format of reviews. As pointed out by our cross-cultural study (*H1*), online consumers in China are not self-motivated to provide online reviews. The number of reviews provided by Chinese consumers for each book is much smaller than that by consumers in the US (see [Table 2](#)). Chinese are likely to provide much shorter reviews than consumers in the US (*H3* of the first study), and by examining the review data collected from Amazon China, we find that consumer in China tend to provide meaningless reviews. For example, most of the reviews in Amazon China are short sentences like "like it" or "dislike it", and express nothing else but similar content as the emotional tendency (stars) of reviews. This can partly explain the result of our second study (*H5*) that why on Dangdang a longer review cannot promote more sales. Therefore, in the Chinese context, reviews with standardized format can not only encourage users to provide useful reviews, but also make it more convenient for readers to get their needed information at low cost. In the premise of free speech environment, reviews should be standardized as best as possible. This can be achieved through conducting related studies on reviews' readability to make sure that reviews will be standardized effectively.

Second, strengthen and promote review voting mechanism. Based on the study of [Chen \(2008\)](#), we know that reviews with a larger number of "helpful" votes can have more significant effect on purchase decision making of potential consumers on Amazon. However, this mechanism does not work on Dangdang because "helpful" votes of reviews are found not to be significantly corre-

lated with book sales (*H4* of the second study). This can be explained partly by our cross-cultural study (*H4*). Only a limited proportion of online reviews has ever received votes from other users in China. Therefore, Chinese online review systems should put more efforts to study behavioral habits, psychological aspect and participating motivation of Chinese people with the purpose of effectively encouraging readers to vote for reviews. Additionally, incentive mechanisms ([Zhang 2009](#)) should be created to encourage voting participation and feedback provision so that review voting mechanism can truly promote product sales. In some cases, especially in the early stage of promoting voting system, some compulsory measures (e.g., where review readers are required to vote the reviews) can be taken in the Chinese context.

Third, further improve spotlight reviews. Our second study (*H6*) shows that spotlight reviews have more positive effect on product sales than other reviews. Compared with consumers in the US, consumers in China are more willing to follow the spotlight reviews (*H5* of the first study), which can also encourage engagement in review systems. Therefore, for online review systems in China, spotlight reviews mechanism should be deepened both from connotation and extension. For example, apart from editor recommendation, review voting mechanism and reviewer rank mechanism both can be used to produce spotlight reviews and then improve the quality of spotlight reviews. These are especially feasible in the Chinese context, because information generated by online consumers can better promote the popularity of products or service than that created by editors ([Zhang et al. 2010](#)). Thus, not only the cost of information search can be reduced, but also consumers' interest in e-commerce will be enhanced, which then promotes product sales. This is applicable both in the Chinese context and the US context.

### 6.4. Activeness of reviewers

Our second study (*H1*) proves that number of reviews is positively correlated with product sales, which is consistent with former empirical studies toward Amazon and other e-markets. Therefore, for e-commerce websites, we should emphasize on encouraging consumers to participate in online communities. This is especially important in the Chinese context since online consumers are not self-motivated to participant in online review systems, and reluctant to provide reviews and votes for others' reviews (*H1, H2 and H4* of the first study). Together with considering reviewer rank and quality of reviews, a Chinese website can differentiate its online review system with those of other e-commerce websites. Only through this way can the Chinese

e-commerce website take a firm place in this e-commerce market featured with low entry threshold and being easily imitated.

In conclusion, the purpose of improving online review system is to enhance the invisible information effect of online reviews. Hence it is necessary to investigate how to effectively maximize the effect of online review systems from both macro and micro perspective. During this investigation, different factors should be considered, including shopping habits, user behavior and Chinese culture. Moreover, standardizing virtual communities and improving the credibility of online review systems should be constantly and well thought by all the Chinese online stores.

## 7. Conclusions and future work

In this paper, we study the problem of how to develop an effective online review system in the Chinese context by conducting two empirical studies.

Firstly, we conduct a cross-cultural study to examine how the national culture difference across countries influences online consumers' behavior of providing reviews. We propose our hypotheses for different aspects of reviews based on Hofstede's framework of national culture. The aspects of reviews include number of reviews, emotional tendency of reviews, length of reviews, spotlight reviews, and the number of helpful and helpless votes. Cultural dimensions in the framework consist of power distance, masculinity *versus* femininity, individualism *versus* collectivism, uncertainty avoidance and long-term orientation. We test our proposed hypotheses based on the analysis of consumer reviews on books collected from Amazon US and Amazon China respectively. Our results show that: (1) online consumers are influenced by culture difference across countries when providing online reviews, but the effect of power distance with regard to spotlight reviews will become smaller when corresponding books become more popular online; and (2) Chinese consumers are reluctant to engage in the online review systems and "helpfulness" voting mechanism compared with their American counterparts, but they tend to hold more positive attitude towards products or service, provide less extremely negative reviews, and weight more on the negative reviews provided by other online consumers.

Secondly, we investigate the impact of online reviews on product sales in the Chinese context. We first propose our hypotheses about the correlations between different aspects of online reviews and product sales based on previous theories and research findings. The constructs related to reviews include number of reviews, reputation of reviewers, emotional tendency of reviews, quality of reviews and timeliness of reviews. To test our hypotheses, we propose a linear panel econometric model and fit the model in our crawled dataset. The dataset is crawled from Dangdang, which is the largest online book store in China. Our results show that: (1) number of reviews, emotional tendency of reviews and spotlight reviews are all positively correlated with product sales, while reviewer rank and length of reviews have negative effect on product sales; and (2) the review voting mechanism of Dangdang is not significantly correlated with product sales.

Based on our two empirical studies and the findings in others' studies, we propose several guiding suggestions for online review systems in the Chinese context. Firstly, online review systems should rank reviewers in a more reasonable manner so that consumers could express their emotional tendency, pay attention to negative reviews, and standardize the format of reviews. In addition, designers should enhance the effectiveness of review voting mechanisms and improve the quality of spotlight reviews so that online consumers can reduce their information search cost and find their needed information quickly. Furthermore, online review systems should encourage consumers to share more real experience

on purchase and usage. The truthfulness of reviews should also be taken into account for reviewer rank evaluation, review voting mechanism and spotlight reviewers selection.

The contributions of our current work are as follows: (1) we explore the impact of national culture difference across countries (the US and China in this paper) on online consumers' behavior of providing reviews by analyzing the real data collected from Amazon US and Amazon China. Different from other cross-cultural studies, we choose our two datasets from two websites belonging to the same corporation but running in the US and China respectively. This will partly eliminate biases caused by differences of websites for our cross-culture study; (2) different from previous studies on the exploration of influence of online reviews on product sales, we use the amount of product sales rather than the rank of product sale as dependent variable in our model. Since the rank data is not as suitable as continuous amount data for the linear model, we believe that our model could lead to more reliable results; (3) our model could capture the dynamic impact of online reviews on product sales in the Chinese context, since we use the panel data and model; and (4) we design a comprehensive framework of relationships between online reviews and product sales. We also provide the guiding suggestions for the design of online review systems in the Chinese context.

The limitations and future work are discussed as follows: (1) our cross-cultural study only considers the numeric information of aspects of online reviews and ignores the qualitative information hidden in text segments. Consequently, our hypotheses only cover part of the connections between online reviews and dimensions of national culture. For future work, we will employ text mining techniques to discover hidden information in text segments and take into account that information in the model; (2) our cross-culture study only focuses on one particular product – book. In the future, we will investigate effects of properties of different products. Specifically, we could conduct studies for products in different categories, or for products with different levels of popularity; and (3) in our second study, our econometric model does not consider the lagged effect of reviews on sales and the endogenous effect of variables. Besides, the comparison between our results on Dangdang and others' findings on Amazon US is not robust, since these studies might be conducted in different conditions. In the future work, we will try to eliminate the bias caused by the different conditions by designing a more refined comparative study.

## References

- Ba, S., and Pavlou, P. A. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *MIS Quarterly*, 26, 2002, 243–268.
- Bickart, B., and Schindler, R. M. Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15, 2001, 31–40.
- Bricks, D. A. *Bunders in International Business*, 4th edition. Blackwell, 2006.
- Brynjolfsson, E., Hu, Y., and Smith, M. D. Consumer surplus in the digital economy: Estimating the value of increased product variety. *Management Science*, 2003, 49.
- Chen, B. L. Empirical Study of Online Word-of-Mouth: Based on Network and Trust Theory. PhD dissertation, Zhe Jiang University, China, 2008.
- Chen, P. Y., Dhansabhorn, S. and Smith, M. D. All reviews are not created equal: The disaggregated impact of reviews and reviewers at amazon.com. In Proceeding of the 28th Annual International Conference on Information Systems (ICIS 2007), 2007.
- Chen, P. Y., Wu, S. Y. and Yoon, J. S. The impact of online recommendations and consumer feedback on sales. In Proceedings of the 25th International Conference on Information Systems (ICIS), 2004.
- Chevalier, J., and Mayzlin, D. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 8, 2006, 345–354.
- China Internet Network Information Center (CNNIC). The 29th Statistical Report on Internet Development in China, 2012.
- comScore. Amazon.com drew over 282m unique visitors in June. <http://finance.yahoo.com/news/comScore-Amazon-com-drew-282M-theflyonthewall-2098098523.html>. Accessed July 18, 2012.
- Dellarocas, C. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, 49, 2005, 1407–1424.



- Dellarocas, C., Zhang, X. M., and Awad, N. F. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, 21, 2007, 23–45.
- Doh, S. J., and Hwang, J. S. How consumers evaluate ewom (electronic word-of-mouth) messages. *Cyber Psychology & Behavior*, 12, 2005, 56–72.
- Dorfman, P. W., and Howell, J. P. Dimensions of national culture and effective leadership patterns: Hofstede revisited. *Advances in International Comparative Management*, 3, 1988, 127–150.
- Duan, W. J., Gu, B., and Andrew, D. W. Do online reviews matter? – An empirical investigation of panel data. *Decision Support Systems*, 2008, 1007–1016.
- Duan, W. J., Gu, B., and Whinston, A. B. The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *Journal of Retailing*, 84, 2008, 233–242.
- Fang, H., Zhu, Q. H. and Zhang, J. An empirical analysis of the impact of online reviews on product sales in the Chinese context. In Proceedings of the 2011 IEEE International Conference on Advanced Information Networking and Applications (AINA), 2011, 43–54.
- Fernandez, D. R., Carlson, D. S., Stepina, L. P., and Nicholson, J. D. Hofstede's country classification 25 years later. *The Journal of Social Psychology*, 137, 1997, 43–54.
- Ghose, A., and Ipeiritos, P. G. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics, 2010. <http://ssrn.com/abstract=1261751>.
- Godes, D., and Mayzlin, D. Using online conversations to study word-of-mouth communication. *Marketing Science*, 23, 2004, 545–560.
- Guadagno, R. E., and Cialdini, R. B. Online persuasion and compliance: Social influence on the Internet and beyond. In Y. Amichai-Hamburger (ed.), *The Social Net: The Social Psychology of the Internet*, 2003.
- Hao, Y. Y., Li, Y. J. and Ye, Q. Dynamic impacts of online reviews and other information sources on sales in panel data environment: Evidence from movie industry. In Proceedings of the International Conference on Management Science & Engineering, 2008.
- Hennig-Thurau, T., Gwinner, K. P., and Walsh, G. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet. *Journal of Interactive Marketing*, 18, 2004, 38–52.
- Hofstede, G. *Culture's Consequence*. Sage, Beverly Hills, 1980.
- Hofstede, G. *Culture's Consequence*, 2nd edition. Sage, Thousand Oaks, 2001.
- Hofstede, G., and Bond, M. H. Hofstede's culture dimensions: An independent validation using Rokeach's value survey. *Journal of Cross-Cultural Psychology*, 15, 1984, 417–433.
- Hofstede, G., and Hofstede, G. J. Dimensions of national cultures. <http://www.geerthofstede.nl/culture/dimensions-of-national-cultures.aspx>. Accessed March, 2012.
- Hofstede, G., Hofstede, G. J., and Minkov, M. *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival*, 3rd edition. McGraw-Hill, 2010.
- Hu, N., Liu, L., and Sambamurthy, V. Fraud detection in online consumer reviews. *Decision Support Systems*, 50, 2011, 614–626.
- Hu, N., Liu, L., and Zhang, J. Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management*, 9, 2008, 201–214.
- Hu, N., Pavlou, P. A. and Zhang, J. Can online reviews reveal a product's true quality? Empirical findings and analytical modeling of online word-of-mouth communication. In Proceedings of the 7th ACM Conference on Electronic Commerce, 2006.
- iResearch. User access ranking top10 Chinese B2C websites, 2011. <http://www.iresearch.com.cn/View/161922.html>. Accessed July 18, 2012.
- Josang, A., Ismail, R., and Boyd, C. A survey of trust and reputation systems for online service provision. *Decision Support Systems*, 43, 2007, 618–644.
- Koh, N. S., Hu, N., and Clemons, E. K. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electronic Commerce Research and Applications*, 9, 2010, 374–385.
- Korfiatis, N., Bariocanal, E. G., and Alonso, S. S. Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11, 2011, 205–217.
- Language, C. How to calculate the number of Chinese words in English-to-Chinese translation services? <http://www.scientrans.com/service-word-account.html>. Accessed March, 2012.
- Lee, J., Park, D., and Han, I. The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, 7, 2008, 341–352.
- Lee, J., Park, D. H., and Han, I. The different effects of online consumer reviews on consumers' purchase intentions depending on trust in online shopping malls: An advertising perspective. *Internet Research*, 21, 2011, 187–206.
- McCrae, R. R., and Terracciano, A. Personality profiles of cultures: Aggregate personality traits. *Journal of Personality and Social Psychology*, 89, 2005, 407–425.
- McSweeney, B. Hofstede's model of national cultural differences and their consequences: A triumph of faith and a failure of analysis. *Human Relations*, 55, 2002, 89–118.
- Merton, R. K. The Matthew effect in science. *Science*, 159, 1968, 56–63.
- Nakata, C., and Sivakumar, K. Instituting the marketing concept in a multinational settings: Implications for a changing culture. *Journal of the Academy of Marketing Science*, 29, 2001, 255–275.
- Park, D. H., and Kim, S. The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*, 7, 2008, 399–400.
- Pavlou, P. A., and Chai, L. What drives electronic commerce across cultures? A cross-cultural empirical investigation of the theory of planned behavior. *Journal of Electronic Commerce Research*, 3, 2002, 240–253.
- Picazo-Vela, S., Chou, S. Y., Melcher, A. J., and Pearson, J. M. Why provide an online review? An extended theory of planned behavior and the role of big-five personality traits. *Computers in Human Behavior*, 26, 2010, 685–696.
- Resnick, P., and Zeckhauser, R. Trust among strangers in internet transactions: Empirical analysis of eBay's reputation system. *The Economics of the Internet and E-Commerce. Advances in Applied Microeconomics*, 11, 2002, 127–157.
- Schnapp, M., and Allwine, T. Mining of book data from Amazon.com, presentation at the ucbsims web mining conference, 2001. <http://www2.sims.berkeley.edu/resources/affiliates/workshops/webmining/Slides/ORA.ppt>.
- Smith, M., and Brynjolfsson, E. Customer decision making at an internet shopbot: Brand still matters. *The Journal of Industrial Economics*, 49, 2001, 541–558.
- Sondergaard, M. Research note: Hofstede's consequence: A study of reviews, citations and replications. *Organization Studies*, 15, 1994, 447–456.
- Tang, L. L. Comparative Study on Chinese and Foreign Online Bookstores: A Research Based on Dangdang and Amazon. PhD thesis, Huazhong Normal University, 2008.
- Utz, S., Kerkhof, P., and Bos, J. V. D. Consumers rule: How consumer reviews influence perceived trustworthiness of online stores. *Electronic Commerce Research and Applications*, 2011. <http://dx.doi.org/10.1016/j.elerap.2011.07.010>.
- Wang, J., Wang, G. G., Ruona, W. E. A., and Rojewski, J. W. Confucian values and the implications for international HRD. *Human Resource Development International*, 8, 2005, 311–326.
- Wang, S. C. *Individual/Organizational Characteristics and Intention to Adopt E-Commerce: A Study based on Innovation Adoption Theory*. The Chinese University of Hong Kong, China, 2001.
- Wikipedia. Amazon China. [http://en.wikipedia.org/wiki/Amazon\\_China](http://en.wikipedia.org/wiki/Amazon_China). Accessed July 18, 2012.
- Ye, Q., Zhang, Z. Q., and Luo, Z. X. Research on Chinese subjectivity automatically analysis towards emotional tendency of online reviews. *China Journal of Information Systems*, 1, 2007, 79–91.
- Zhang, J. Promoting Honesty in Electronic Marketplaces: Combining Trust Modeling and Incentive Mechanism Design. PhD thesis, University of Waterloo, 2009.
- Zhang, Z. Q., Ye, Q., Law, R., and Li, Y. J. The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews. *International Journal of Hospitality Management*, 29, 2010, 694–700.