Marketing activity, blogging and sales

Hiroshi Onishi a, Puneet Manchanda b,*

a Dentsu Inc., Japan
b Ross School of Business at the University of Michigan, USA

1. Introduction

Consumer Generated Media (CGM), such as blogs (a contraction of the term “Web logs”), have witnessed explosive growth in recent years. For example, the number of blogs worldwide was estimated to be 184 million, with a readership of 346 million (Technorati/Universal McCann, March 2008). In contrast, there were virtually no blogs in March 2003. Other types of CGM have also experienced similar growth patterns, such as Facebook, which was launched in February 2004 and currently has approximately 500 million members worldwide (in October 2007). These statistics clearly show that there is considerable CGM activity (e.g., multi-media posting, blogging, visits, traffic) by consumers. Given that CGM (almost always) co-exist with traditional media, an important research question is whether CGM content helps or hurts traditional media content. This distinction is also important from a managerial perspective, as traditional media—in which a manufacturer creates and delivers content to consumers—consume the marketing resources of firms. In contrast with these paid media, “new” media (in which consumers create content and this content is exchanged between other consumers and potentially between manufacturers) are primarily available for free. This question becomes even more salient when new product launches are involved because firms typically allocate approximately half of their marketing budgets to support new products.

One of the most prevalent forms of new media is blogging. Therefore, we assemble a unique data set from Japan that contains market outcomes (sales) for new products, new media (blogs) and traditional media (TV advertising) in the movie category. We specify a simultaneous equation log-linear system for market outcomes and the volume of blogs. Our results suggest that new and traditional media act synergistically, that pre-launch TV advertising spurs blogging activity but becomes less effective during the post-launch period and that market outcomes have an effect on blogging quantity. We find detailed support for some of these results via a unique and novel text-mining analysis and replicate our findings for a second product category, cellular phone service. We also discuss the managerial implications of our findings.
US Internet users) and approximately 8 million in Japan (approximately 5% of all Japanese Internet users) in 2007 (Technorati/Universal McCann, 2008). However, if one examines the total number of posts by language, Japanese-language posts accounted for 37% of all posts worldwide, followed closely by English-language posts at 36% (Technorati, 2007). In addition, the readership of blogs in these two markets is high: approximately half of all Internet users in the US and approximately one-fifth of all Japanese Internet users have read a blog during the past year. There are also indications that blogs are currently considered to be similar to mainstream media sites; the number of blog sites in the top 100 most popular sites (blogs and mainstream media) worldwide was 22 in 2007, and blogs were being viewed by consumers as “sites for news, information, gossip, etc.” (Technorati, 2007). In 2008, four of the top ten entertainment sites were blogs (Technorati/Universal McCann, 2008). Finally, in 2007 (the year in which our data were collected), blogging was likely the most pervasive CGM activity on the web.

Although existing research has investigated the relationship between new media and market outcomes, such studies have generally not accounted for the presence of traditional media. Therefore, the question of whether traditional new media and traditional media are synergistic or antagonistic has not been answered. The evolving nature of this relationship (among market outcomes, new media and traditional media) for a new product before and after its launch has also received scant attention in the literature (see Table 1). Specifically, for blogs (the form of CGM that we consider), there is virtually no empirical research that documents the value of blog volume and quality (including information content) in predicting sales, especially in a setting in which traditional media are also present. This lack of research is not surprising because blogs, which are personal, non-directed communication tools, are unique in the world of CGM. For example, Kumar, Novak, Raghavan, and Tomkins (2005) note that blogs are unique for sociological reasons: they comprise a “highly dynamic, temporal community structure” that “focuses heavily on local community interactions,” and for technical reasons, blogs “offer us a ready-made view of evolution (of content) in continuous time.” A survey (Technorati, 2009) also revealed that 75% of the respondents write blogs for enjoyment and to record their personal musings rather than to evaluate products. Consequently, it is unclear whether blogging activity is related to traditional media and market outcomes.

The findings from the (limited) current research on the effectiveness of new media provide some additional motivation for this research. The first set of findings pertains to the direct relationship between blogs and (implied) market outcomes. The general finding suggests a positive correlation between blogging activity and sales ranks on e-commerce sites, such as Amazon.com. For example, Dhar and Chang (2009) find a positive correlation between music

---

1 Most of the published literature does not consider the relationship between new and traditional media and has typically limited its focus on establishing a link between online chatter (newsgroup postings, reviews and ratings) and its effect on market outcomes. These papers have provided some evidence of a positive correlation between the volume of online user ratings and sales (Chintagunta et al., 2010; Dellarocas et al., 2007; Duan et al., 2008; Karniouchina, 2011; Liu, 2006). Given the goal-driven nature of product ratings (and reviews), these findings may not be unexpected. Other research has looked at the synergies between different forms of traditional media e.g., online and offline (Naik & Peters, 2000).

2 There is also emerging work on other (non-market outcome) aspects of blogs, such as linking decisions between and across blog posts (e.g., Mayzlin & Yoganarasimhan, in press).

3 Although sales ranks are useful indicators of market outcomes, they are limited in their applicability because (a) there is a potential for the rank-sales relationship to be noisy (especially if Amazon.com is not a representative seller of the goods in question), (b) ranking data are idiosyncratic and usually not available for a wide variety of products (e.g., movie box-office sales, packaged goods, services) and (c) it is unclear whether managers can use the imputed sales data to obtain actual effect sizes to set policy.

---

Table 1

Overview of current research on CGM.

<table>
<thead>
<tr>
<th>Product category</th>
<th>Online CGM</th>
<th>New and Traditional Media interaction</th>
<th>CGM measures</th>
<th>Outcome measures</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper</td>
<td>Pre-launch</td>
<td>Sales volume (audience size)</td>
<td>Aggregate</td>
<td>Sales volume, valence, intensity</td>
<td>Daily/monthly panel</td>
</tr>
<tr>
<td>Dhar and Chang (2009)</td>
<td>Blog, Yahoo!</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Karniouchina (2011)</td>
<td>Blog, Unusual McCann</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Gruhl et al. (2005)</td>
<td>Blog, No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Liu (2006)</td>
<td>Blog, Yahoo!</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Moul (2007)</td>
<td>Blog, Yahoo!</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Chevalier and Godes (2004)</td>
<td>Blog, Yahoo!</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Godes and Mayzlin (2004)</td>
<td>Blog, Yahoo!</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Trusov, Bodapati, and Bucklin (2009)</td>
<td>Blog, Yahoo!</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>
album sales (imputed via sales ranks on Amazon.com) and online chatter (as found in blogs and on social networks). Using data for 108 music albums in early 2007, these authors find a positive correlation between future music sales and both the number of blogs and Myspace member activity. Gruhl, Guha, Kumar, Novak, and Tomkins (2005) propose a new methodology to automatically generate a query of blog keywords to detect increases in book sales ranks for Amazon.com, and they report a similar positive correlation. The second set of findings pertains to the relationship between traditional media and word of mouth. Examining multiple product categories, Graham and Havlena (2007) find a positive correlation among advertising, offline word of mouth and online brand comments. However, these authors do not consider market outcomes in their analysis.

Our results suggest that the interaction between blogs and TV advertising has a positive effect on market outcomes. In other words, new and traditional media are synergistic. When we examine the nature of this relationship more carefully with respect to product launch, we find that the relationship becomes weaker after launch. Our explanation for this change is that TV advertising can independently increase pre-launch blogging via the provision of information and content. In contrast, consumers may rely less on traditional media and more on actual product attributes following a product launch; thus, a much weaker relationship between new and old media may emerge. We find support for this “process” explanation via a novel text-mining analysis that we incorporate into our model. To our knowledge, the documentation of this synergistic relationship between traditional and new media as well as the shift in its nature following product launch is a new contribution to the literature.

We quantify the managerial effects of our results using a simulation. The results of the simulation show that a 1% increase in advertising stock results in approximately a ten-per-cent increase in sales in the short term. Interestingly, approximately three-fourths of this increase can be attributed to the synergistic relationship between blogs and advertising; the remainder of the increase reflects the direct effect of advertising. This result suggests that managers should leverage the synergistic relationship between new and traditional media by allocating resources after accounting for the effect of the

Table 2
Movie list.

<table>
<thead>
<tr>
<th>Title</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spider-Man 3</td>
<td>US</td>
</tr>
<tr>
<td>300</td>
<td>US</td>
</tr>
<tr>
<td>Apocalypto</td>
<td>US</td>
</tr>
<tr>
<td>Die Hard 4.0</td>
<td>US</td>
</tr>
<tr>
<td>The Prestige</td>
<td>US and British</td>
</tr>
<tr>
<td>VolVER</td>
<td>European</td>
</tr>
<tr>
<td>The Magic Flute</td>
<td>European</td>
</tr>
<tr>
<td>Confession of Pain</td>
<td>Hong Kong</td>
</tr>
<tr>
<td>Maiko Haaaan!!!</td>
<td>Japanese</td>
</tr>
<tr>
<td>Sayuki</td>
<td>Japanese</td>
</tr>
<tr>
<td>Dai Nipponjin</td>
<td>Japanese</td>
</tr>
<tr>
<td>Shrek the Third</td>
<td>Animation</td>
</tr>
</tbody>
</table>

Table 4
Traditional media (TV GRPs).

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All periods</td>
<td>3047.60</td>
<td>9621.23</td>
</tr>
<tr>
<td>Pre-launch</td>
<td>3471.31</td>
<td>10,553.11</td>
</tr>
<tr>
<td>Post-launch</td>
<td>1992.98</td>
<td>6657.42</td>
</tr>
</tbody>
</table>

spillover from traditional to new media. We also find that the use of new media assists in predicting sales.

The remainder of the paper is organized as follows. Section 2 describes the institutional setting, the data and the text-mining procedure. We discuss the model in Section 3. Section 4 contains the results and the managerial implications. We replicate our findings with data from another product category in Section 5. We conclude the paper in Section 6.

2. Institutional setting and data

Our data are obtained from the Japanese market. As mentioned previously, Japan is the second-largest market worldwide in blog participation and in the number of blog posts. Our main analysis focuses on data from movie launches in Japan. In a subsequent section, we replicate our results using data from the cellular phone service category. We first describe the market outcome data and then describe the measurement of traditional and new media.

2.1. Market outcomes

The outcome variable that we use for movies is the size of a movie’s audience each day after the launch of the movie. These data were obtained from Kyogyo Tsushinsha, an industry periodical. We have data on twelve major movies that were released (launched) during the period from January 2007 to August 2007. The movie titles were chosen to reflect sufficient variation in movie genres (e.g., action, animation) among American, European and Japanese movies. The list of movie titles is presented in Table 2, and Table 3 describes the daily post-launch sales patterns.

2.2. Traditional media

The traditional marketing variable that we use is TV advertising, which was measured in units of daily Gross Rating Points (GRPs). Dentsu Inc. (the largest advertising agency in Japan) provided us with the measure of national GRPs delivered to Japanese households for the movies in our data set. Table 4 shows some differences in pre- and post-launch TV advertising patterns. Generally, the pre-launch TV GRPs were larger (on average) than the post-launch TV GRPs. Unsurprisingly, peak advertising for a movie occurred a week before its launch date to generate high demand at the time of the movie opening.

2.3. New media

We obtain blogging data from Dentsu Buzz Research Inc. (www.dbuzz.jp). This company scans and indexes the major blogging sites in Japan on a daily basis using keywords that cover approximately 64% of all blog articles in 2008. These data are then aggregated, and the count of the daily number of blogs that mention a particular keyword on a given day (multiple mentions on the same day on the same blog are counted as a single mention) is provided to us. A typical blog based on the launch of the movie “Spider-Man 3” is shown in Fig. 1. As is typical for most blogs, the blog contents appear in reverse

---

Table 3
Market outcomes.

<table>
<thead>
<tr>
<th>Description</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily audience</td>
<td>1298.95</td>
<td>4716.47</td>
</tr>
</tbody>
</table>
chronological order and include the blogger’s profile, “trackbacks” (links showing other websites, typically other blogs, to which a blog is linked) and comments. For our product category (movies), we have access to blogs that contain movie mentions from the beginning (i.e., we do not encounter the problem of left truncation).

Buzz Research archives the contents of all blog posts; conducts a lexical analysis of the contents of each tracked blog using proprietary text-mining methods; and classifies the valence of each blog entry as positive, negative or neutral (because this procedure is proprietary, we do not have any information regarding the actual methodology that is used to generate this classification). We have access to the actual content of all posts and the daily number of positive, negative and neutral blogs for the movie and cellular phone service markets. Tables 5 and 6 present details regarding the percentages of blogs pre- and post-launch by valence (positive, negative and neutral). As shown in the tables, the average number of blogs per period increased substantially following the movie launches. Interestingly, the greatest relative growth was observed for neutral blogs.

To illustrate the relationship between marketing outcomes and both traditional and new media, we select a product from among our three product markets. Fig. 2 contains the data series for the movie “Spider-Man 3.” The figure suggests that TV advertising, blog volume and movie viewership were temporally correlated. Dividing the data temporally at the date of launch (May 1, 2007), we observe that the TV GRPs and the number of blogs exhibit an increasing trend pre-launch but a decreasing trend post-launch. Although we illustrate a typical data pattern through this example, note that the pattern was not identical for all brands.

2.4. Text mining

As we have access to all blog posts, we supplement the above blog quantity and valence data with summary results from a text-mining analysis of these posts. We expect that the use of text mining will enable us to obtain some process insights that would be otherwise difficult to obtain using the aggregate data described above.

We were unable to find a standard approach for text-mining Japanese-language text in the literature. Therefore, we based our

<table>
<thead>
<tr>
<th>Table 5</th>
<th>New media (number of blogs).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period</td>
<td>Mean</td>
</tr>
<tr>
<td>All periods</td>
<td>99.43</td>
</tr>
<tr>
<td>Pre-launch</td>
<td>43.78</td>
</tr>
<tr>
<td>Post-launch</td>
<td>237.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Blog valence percentage.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence of blogs</td>
<td>Period</td>
</tr>
<tr>
<td>Positive blogs</td>
<td>All periods</td>
</tr>
<tr>
<td></td>
<td>Pre-launch</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
</tr>
<tr>
<td>Neutral blogs</td>
<td>All periods</td>
</tr>
<tr>
<td></td>
<td>Pre-launch</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
</tr>
<tr>
<td>Negative blogs</td>
<td>All periods</td>
</tr>
<tr>
<td></td>
<td>Pre-launch</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
</tr>
</tbody>
</table>
approach on several other studies that have performed text-mining on Japanese text (Matsumura, Yamamoto, & Tomozawa, 2008; Yamamoto & Matsumura, 2009). Our procedure was as follows. First, we examined all posts across all blogs, which totaled approximately 200,000. Using a lexical software program (KH Coder) for Japanese text, we then performed the following steps:

1. We parsed each sentence in the blog posts into words. This step was important because Japanese-language text is not naturally delimited into words. We then classified the words into lexical categories (e.g., nouns, verbs).
2. We eliminated all proper nouns, articles and prepositions.
3. We assessed the words for tense (e.g., the same word in three different tenses is counted as three unique words).
4. We performed a frequency count of all unique words.
5. We selected the words that were relevant to our analysis.
6. We computed the number of times that a relevant word appeared across all blog posts for the temporal unit of analysis (e.g., day or week).

This task was laborious because it involved frequent human intervention to ensure that the parsing, classification and word frequencies were correct. We performed this process for each of the blog posts. The top 30 words are listed according to their frequency in Table 7. We also show the frequencies for words that were germane to our analysis, such as “advertising” (see discussion below).

As noted above, our objective was to provide process explanations for our findings. We conducted exploratory analyses using all 30 words and finally retained the words that were relevant and appeared to have general predictive power. The retained words were “Advertising,” “Award,” "Interesting" and “Viewed.” We conjectured that these words were found to have some explanatory power because consumers would find posts that contain these words (especially “Award,” "Interesting" and “Viewed”) to be useful in terms of making their own decisions with regard to viewing the movies. We also expected “Advertising” and “Award” to represent a higher proportion of blog posts pre-launch because they reflected

<table>
<thead>
<tr>
<th>Word</th>
<th>Rank</th>
<th>Frequency</th>
<th>Lexical category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Movie</td>
<td>1</td>
<td>19,362</td>
<td>Noun</td>
</tr>
<tr>
<td>Dainipponjin</td>
<td>2</td>
<td>10,284</td>
<td>Proper noun</td>
</tr>
<tr>
<td>View</td>
<td>3</td>
<td>16,150</td>
<td>Verb</td>
</tr>
<tr>
<td>Matsumoto Hitoshi</td>
<td>4</td>
<td>9259</td>
<td>Proper noun</td>
</tr>
<tr>
<td>Think</td>
<td>5</td>
<td>8508</td>
<td>Verb</td>
</tr>
<tr>
<td>Director</td>
<td>6</td>
<td>8402</td>
<td>Noun</td>
</tr>
<tr>
<td>Go</td>
<td>7</td>
<td>4701</td>
<td>Verb</td>
</tr>
<tr>
<td>Say</td>
<td>8</td>
<td>4611</td>
<td>Verb</td>
</tr>
<tr>
<td>Interesting</td>
<td>9</td>
<td>4372</td>
<td>Adjective</td>
</tr>
<tr>
<td>Film</td>
<td>10</td>
<td>4,281</td>
<td>Noun</td>
</tr>
<tr>
<td>Award</td>
<td>11</td>
<td>3,228</td>
<td>Noun</td>
</tr>
<tr>
<td>Release</td>
<td>12</td>
<td>3,152</td>
<td>Verb</td>
</tr>
<tr>
<td>Warai</td>
<td>13</td>
<td>2,723</td>
<td>Proper noun</td>
</tr>
<tr>
<td>Feeling</td>
<td>14</td>
<td>2,677</td>
<td>Noun</td>
</tr>
<tr>
<td>Theater</td>
<td>15</td>
<td>2,509</td>
<td>Noun</td>
</tr>
<tr>
<td>Out</td>
<td>16</td>
<td>2,380</td>
<td>Verb</td>
</tr>
<tr>
<td>Story</td>
<td>17</td>
<td>2,352</td>
<td>Noun</td>
</tr>
<tr>
<td>Japan</td>
<td>18</td>
<td>2,287</td>
<td>Noun</td>
</tr>
<tr>
<td>Fun</td>
<td>19</td>
<td>2,174</td>
<td>Noun</td>
</tr>
<tr>
<td>Love</td>
<td>20</td>
<td>2,002</td>
<td>Verb</td>
</tr>
<tr>
<td>Self</td>
<td>21</td>
<td>1,900</td>
<td>Noun</td>
</tr>
<tr>
<td>Interview</td>
<td>22</td>
<td>1,891</td>
<td>Noun</td>
</tr>
<tr>
<td>News</td>
<td>23</td>
<td>1,882</td>
<td>Noun</td>
</tr>
<tr>
<td>Content</td>
<td>24</td>
<td>1,825</td>
<td>Noun</td>
</tr>
<tr>
<td>TV</td>
<td>25</td>
<td>1,811</td>
<td>Noun</td>
</tr>
<tr>
<td>Matsu</td>
<td>26</td>
<td>1,782</td>
<td>Proper noun</td>
</tr>
<tr>
<td>Down Town</td>
<td>27</td>
<td>1,733</td>
<td>Proper noun</td>
</tr>
<tr>
<td>Laugh</td>
<td>28</td>
<td>1,718</td>
<td>Verb</td>
</tr>
<tr>
<td>Kitano Takeshi</td>
<td>29</td>
<td>1,603</td>
<td>Proper noun</td>
</tr>
<tr>
<td>Write</td>
<td>30</td>
<td>1,586</td>
<td>Verb</td>
</tr>
<tr>
<td>Advertising</td>
<td>80</td>
<td>686</td>
<td>Noun</td>
</tr>
</tbody>
</table>

The discussion of awards may seem surprising for movies that had only recently been released. All of the non-Japanese movies in our sample had previously been released in their home countries. Thus, these movies had been candidates for awards or, in some cases, had already won awards.
informational attributes that were available pre-launch, and we expected “Viewed” to represent a higher proportion of posts post-launch and to represent discussions based on the movie viewing experiences of consumers. However, “Interesting” could pertain to movies in either the pre- or post-launch periods; therefore, we could not predict the proportion of blog posts that used this word during the pre- or post-launch periods. Table 8 shows the proportion of blogs that used each of the chosen words during the pre- and post-launch periods. As shown in the table, the pre- and post-launch proportions are generally consistent with our expectations. Specifically, bloggers used “Advertising” in their blog posts much more frequently in the pre-launch period and used “Viewing” much more frequently after launch. This finding suggests that traditional advertising is likely used to enrich blog posts pre-launch, whereas service experience and product attribute knowledge is much more important to bloggers (and presumably to their readers) following launch.

In conclusion, these data are novel in that they combine marketing data for both traditional and new media with market outcomes from a market in which new media have proven to be important (at least in terms of activity). These data are also novel in that they enable us to focus on new product launches. The availability of the actual blog post text reveals opportunities for deeper analysis. Finally, we use data from a completely unrelated product category (i.e., cellular phone service) to confirm the robustness and generalizability of our results (Section 5).

### 3. Model

Our model contains two dependent variables: market outcomes (daily movie viewers) and the volume of new media (the number of blogs). As mentioned previously, it is feasible that market outcomes and the volume of new media are determined simultaneously. Therefore, our model takes the form of a system of simultaneous equations (with correlated errors) to account for this potential simultaneity.

In terms of our specification, we consider a range of factors that could drive movie sales (the number of customers). First, we allow movie sales to be a function of sales in the last time period and the cumulative sales (or customers) until the last time period.

Next, for new products, of which the quality is difficult to evaluate prior to purchase, consumers seek information and obtain opinions from other consumers (Harrison-Walker, 2001; Rogers, 1983). One such source of information and opinions is the blogosphere (Dhar & Chang, 2009). Blogs are unique relative to other CGM in terms of activity. These data are also novel in that they enable us to term the robustness and generalizability of our results (Section 5).

### Table 8

<table>
<thead>
<tr>
<th>Word</th>
<th>Period</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viewed</td>
<td>All periods</td>
<td>54.2%</td>
<td>27.1%</td>
</tr>
<tr>
<td></td>
<td>Pre-launch</td>
<td>43.3%</td>
<td>30.7%</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
<td>64.8%</td>
<td>10.3%</td>
</tr>
<tr>
<td></td>
<td>All periods</td>
<td>21.9%</td>
<td>28.4%</td>
</tr>
<tr>
<td>Award</td>
<td>Pre-launch</td>
<td>22.8%</td>
<td>30.8%</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
<td>19.9%</td>
<td>21.9%</td>
</tr>
<tr>
<td></td>
<td>All periods</td>
<td>20.3%</td>
<td>19.3%</td>
</tr>
<tr>
<td>Interesting</td>
<td>Pre-launch</td>
<td>17.6%</td>
<td>21.2%</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
<td>26.4%</td>
<td>11.8%</td>
</tr>
<tr>
<td></td>
<td>All periods</td>
<td>14.0%</td>
<td>26.5%</td>
</tr>
<tr>
<td>Advertising</td>
<td>Pre-launch</td>
<td>15.0%</td>
<td>27.8%</td>
</tr>
<tr>
<td></td>
<td>Post-launch</td>
<td>11.7%</td>
<td>23.0%</td>
</tr>
</tbody>
</table>

### Table 9

<table>
<thead>
<tr>
<th>Word</th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day of the week</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fri</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sat</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sun</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Holiday</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movie1</td>
<td>0.26</td>
<td>0.10</td>
<td>2.57</td>
</tr>
<tr>
<td>Movie2</td>
<td>0.39</td>
<td>0.10</td>
<td>3.88</td>
</tr>
<tr>
<td>Movie3</td>
<td>0.42</td>
<td>0.07</td>
<td>6.06</td>
</tr>
<tr>
<td>Movie4</td>
<td>-0.07</td>
<td>0.08</td>
<td>-0.80</td>
</tr>
<tr>
<td>Movie5</td>
<td>0.45</td>
<td>0.11</td>
<td>4.18</td>
</tr>
<tr>
<td>Movie6</td>
<td>3.06</td>
<td>0.07</td>
<td>4.04</td>
</tr>
<tr>
<td>Movie7</td>
<td>0.14</td>
<td>0.06</td>
<td>2.25</td>
</tr>
<tr>
<td>Movie8</td>
<td>0.28</td>
<td>0.06</td>
<td>4.52</td>
</tr>
<tr>
<td>Movie9</td>
<td>-0.26</td>
<td>0.10</td>
<td>-2.58</td>
</tr>
<tr>
<td>Movie10</td>
<td>0.31</td>
<td>0.05</td>
<td>6.29</td>
</tr>
<tr>
<td>Movie11</td>
<td>-0.22</td>
<td>0.07</td>
<td>-3.11</td>
</tr>
<tr>
<td>#Audience(t−1)</td>
<td>0.28</td>
<td>0.11</td>
<td>2.63</td>
</tr>
<tr>
<td>#Blog(t)</td>
<td>-0.03</td>
<td>0.03</td>
<td>-1.00</td>
</tr>
<tr>
<td>#Blog Positive(t)</td>
<td>0.01</td>
<td>2.36–03</td>
<td>4.28 ***</td>
</tr>
<tr>
<td>#Blog Negative(t)</td>
<td>1.00–03</td>
<td>2.18–03</td>
<td>0.47</td>
</tr>
<tr>
<td>#Blog Viewed(t)</td>
<td>0.30</td>
<td>0.04</td>
<td>7.57 ***</td>
</tr>
<tr>
<td>#Blog Award(t)</td>
<td>-0.02</td>
<td>0.01</td>
<td>-2.20</td>
</tr>
<tr>
<td>#Blog Interesting(t)</td>
<td>-0.01</td>
<td>0.01</td>
<td>-1.39</td>
</tr>
<tr>
<td>#Blog Advertising(t)</td>
<td>-0.01</td>
<td>0.67–03</td>
<td>-1.50</td>
</tr>
<tr>
<td>Blogstock</td>
<td>0.40</td>
<td>0.02</td>
<td>21.02</td>
</tr>
<tr>
<td>Salesstock</td>
<td>0.26</td>
<td>3.52–03</td>
<td>74.03</td>
</tr>
<tr>
<td>Adstock</td>
<td>0.08</td>
<td>0.02</td>
<td>4.50</td>
</tr>
<tr>
<td>Blogstock-Adstock</td>
<td>0.01</td>
<td>2.36–04</td>
<td>44.21 ***</td>
</tr>
</tbody>
</table>

N=1729.

Two-sided significance test symbols: ***p<0.001, **p<0.01, *p<0.05, and ^p<0.10.

#Blog Positive and Negative are relative to #Blog neutral.

Stock variables are defined as in Eq. (4).

Based on the above expectation and the previous research findings of an interaction effect on sales for two types of traditional media (e.g., Naik & Raman, 2003), consumers are likely to not only find information from other consumers and from firms to be directly useful but also find that the two sources of information (and persuasion) could potentially reinforce (or contradict) one another. Therefore, we include an a positive post-launch effect on market outcomes. In addition to the volume of information that is available from others, previous research has also shown that the valence of this information influences evaluations of product performance (Nam, Manchanda, & Chintagunta, 2010;Richins, 1983), especially when both positive and negative information is presented (Dellarocas, Zhang, & Awad, 2007; Rozin & Royzman, 2001). Thus, we also expect the valence of current blogs (the number of positive blogs and the number of negative blogs) to affect market outcomes beyond the volume of blogs. In addition to seeking the opinions of other people, consumers may also respond to the marketing actions of a firm, primarily communication in the form of TV advertising. The relationship between sales (or shares) and advertising is well documented in the marketing (e.g., Lodish et al., 1995) and economics (e.g., Bagwell, 2007) literature. Therefore, we expect to observe a direct effect of current and past advertising on market outcomes.

Based on the above expectation and the previous research findings of an interaction effect on sales for two types of traditional media (e.g., Naik & Raman, 2003), consumers are likely to not only find information from other consumers and from firms to be directly useful but also find that the two sources of information (and persuasion) could potentially reinforce (or contradict) one another. Therefore, we include an

---

6 We also estimated a model with a weekly sales lag and found no material change in our results.
interaction term between the stock of blogging volume (based on current and past blogging volume) and the stock of TV advertising (based on current and past TV advertising). To the best of our knowledge, this is the first attempt to study this interaction; thus, there is no clear theory that predicts the sign of these terms.

To obtain a better understanding of the process by which blogs affect market outcomes, we also include the number of blogs that mention the four retained words (as described in the data section) each day as covariates. We expect that the temporal variation in the use of these words, especially during the pre- and post-launch periods, will provide a richer description of how blogs influence readers and their behavior.

In addition to these variables, we include a set of control variables that are temporal (time trend from launch and day-of-week fixed effects) and cross-sectional (movie-specific fixed effects). These variables control for diffusion patterns that are purely driven by temporal phenomena and movie heterogeneity. For example, the movie fixed effect controls for all (non-varying) idiosyncratic aspects of a movie, such as its stars, director, awards, genre and MPAA rating (all parts that contribute to its overall quality). Other factors that could affect the size of the daily audience (e.g., distribution represented by the number of movie screens) are also partially controlled for by the lagged audience variable and the movie fixed effects.

Therefore, our specification includes lagged sales, cumulative lagged sales, current and past TV advertising, current and past blog activity (volume, valence and content), interaction terms between TV advertising and blog activity, the movie fixed effect and the time-trend and day-of-week fixed effects. The market outcome equation is shown below (Eq. (1)). In this equation and subsequent equations, \(j\) represents a movie title, whereas \(t\) represents time (day).

\[
\ln(\text{#Customers}_t) = \alpha_0 + \alpha_1 \text{Trend}_t + \alpha_2 \text{Season}_t + \alpha_3 \text{Brand}_t + \alpha_4 \ln(\text{#PositiveBlogs}_t) + \alpha_5 \ln(\text{#NegativeBlogs}_t) + \alpha_6 \ln(\text{#ContentsBlogs}_t) + \alpha_7 \ln(\text{#Blogstock}_t) + \alpha_8 \ln(\text{#Adstock}_t) + \alpha_9 \ln(\text{#Blogs}_t) + \alpha_{10} \ln(\text{#Blogstock}_t) - \alpha_{11} \ln(\text{#Adstock}_t) + \epsilon_{1t}.
\]

Two issues regarding this specification need to be addressed. First, advertising may be endogenous. In other words, the level of advertising is related to the error term as a result of some factors that are not observed by the researchers (but are observed by the manager). When we discussed the possibility that advertising may be established based on some unobserved (to us) rule with the advertising agency that provided us the data, we were informed that, advertising plans for movie launches (level of GPRs and schedules) were established far in advance of the launch and were usually not changed after launch. Therefore, this institutional feature of the setting enables us to assume that advertising is exogenous. To ensure that this feature was reflected in our data, we also estimated an alternate version of our model, in which we used advertising as an additional dependent variable. In other words, we estimated a system of equations with three dependent variables (market outcomes, the volume of blogs and advertising GPRs) and allowed the error for each equation to be correlated with the error for the other two equations. We find that our main results are unchanged (details are provided in Appendix A). Because we focus on market outcomes and new media (and because of the above two reasons), we present the simultaneous equation model with two dependent variables (market outcomes and the volume of blogs) as our main model. Second, we do not include price in the specification because prices did not vary during the time period of our data.

We now discuss our second dependent variable. As noted previously, this variable is the count of the number of blogs that mention a particular keyword (movie title) in a specific temporal period, such as a day or a month (multiple mentions in the same temporal unit on the same blog are counted as a single mention).

We hypothesize that as the number of customers who consume a product increases, these customers are more likely to disseminate information (Rogers, 1983). These customers could also directly influence other customers who are their “followers,” as proposed by, for example, Christakis and Fowler (2007) and Nair, Manchanda, and...
Bhatia (2010). Thus, we include the number of current and past customers (the current number of customers and the number of cumulative customers up to the previous period) as independent variables. Second, as noted previously, consumers (including bloggers) cannot test products prior to launch; however, they are exposed to the marketing efforts of the firms (TV advertising in this context). A long history of marketing research has documented that consumers both extract information and form attitudes towards advertised products as well as form attitudes toward the advertisement itself. Attitudes toward advertisements have been found to mediate the relationship between beliefs and attitudes toward brands (Brown & Stayman, 1992; Mitchell & Olson, 1981), which in turn influence purchase intentions. Therefore, we expect bloggers to be influenced by advertising via the development of attitudes with regard to both the ads and the brands and to disseminate opinions regarding these advertisements and brands in their posts. Thus, we include current and past advertising as an explanatory variable. Relative to the pre-launch period, however, we expect that bloggers, especially influential bloggers, will have greater access to actual products and will thus base their posts less on advertising content and more on their own experience with such products and product attributes. To obtain a better understanding of this phenomenon, we also include the four retained words from the text-mining exercise described above as covariates.

Third, to control for the “viral” nature of CGM and the unique information network of blogs (Kumar et al., 2005), we also include the number of past blogs (the number of blogs in the previous time period and the stock of blogs up to the previous time period) as explanatory variables. The stock of blog posts and advertising might act synergistically and lead to a greater number of blog posts. Therefore, we include the interaction term between the blog stock and advertising stock.

Finally, for any simultaneous equation system, we need some exclusion restrictions for identification. Identification can be achieved in the current context, for example, when some independent variable(s) affect the number of blogs but not the market outcomes. The excluded variable that we use is the sum of blog posts for all customers (the current number of customers and the number of cumulative customers up to the previous period) as independent variables. Therefore, this variable is used as a covariate in the blog equation but is not used in the sales equation.

As noted above, we expect the actual launch of a product to represent a structural change in the environment. Therefore, we specify different coefficients for the pre- and post-launch periods for each of the independent variables. Finally, as shown in Eq. (1), we also include the control variables. The resulting specification is shown in Eq. (2).

\[
\begin{align*}
\ln(#Blogs_t) &= \beta_0 + \beta_1 \text{Trend}_t + \beta_2 \text{Season}_t + \beta_3 \text{Brand}_t + \beta_4 \ln(\#\text{preBlogs}_{t-1}) \\
&+ \beta_5 \ln(\#\text{prePositiveBlogs}_{t-1}) + \beta_6 \ln(\#\text{preNegativeBlogs}_{t-1}) \\
&+ \beta_7 \ln(\#\text{preContentBlogs}_{t-1}) + \beta_8 \ln(\#\text{preBlogstock}_{t-1}) \\
&+ \beta_9 \ln(\#\text{preAdstock}_{t-1}) + \beta_{10} \ln(\#\text{preAdstock}_{t-1}) \times \ln(\#\text{preAdstock}_{t-1}) \\
&+ \beta_{11} \ln(\#\text{postCustomers}_{t-1}) + \beta_{12} \ln(\#\text{postCumCustomers}_{t-1}) \\
&+ \beta_{13} \ln(\#\text{postBlogstock}_{t-1}) + \beta_{14} \ln(\#\text{postAdstock}_{t-1}) \\
&+ \beta_{15} \ln(\#\text{postNegativeBlogs}_{t-1}) + \beta_{16} \ln(\#\text{postContentBlogs}_{t-1}) \\
&+ \beta_{17} \ln(\#\text{postBlogstock}_{t-1}) + \beta_{18} \ln(\#\text{postAdstock}_{t-1}) + \epsilon_{2t}.
\end{align*}
\]

As is typical for these models, the errors capture all factors that are not observed by the researcher, such as offline word of mouth. These factors could affect each equation independently and/or jointly, thus leading to simultaneity (as discussed previously). Therefore, we allow the errors in Eqs. (1) and (2) to be jointly distributed bivariate normal as in Eq. (3).

\[
\begin{bmatrix}
\epsilon_{1t} \\
\epsilon_{2t}
\end{bmatrix} \sim \text{MVN}(0, \Sigma), \\
\Sigma = \begin{bmatrix}
\sigma_1^2 & \rho_{\epsilon_1\epsilon_2} \\
\rho_{\epsilon_1\epsilon_2} & \sigma_2^2
\end{bmatrix}
\]

Note that all the stocks (cumulative variables) are defined using the standard distributed lag structure (Rao & Miller, 1975) as shown below in Eq. (4).

\[
\begin{align*}
\text{Salestock}_t &= \gamma_1 \text{Salestock}_{t-1} + \gamma_2 \text{Salestock}_{t-2} \\
\text{Blogstock}_t &= \gamma_3 \text{Blogstock}_{t-1} + \gamma_4 \text{Blogstock}_{t-2} \\
\text{Adstock}_t &= \gamma_5 \text{Adstock}_{t-1} + \gamma_6 \text{Adstock}_{t-2}
\end{align*}
\]

In terms of estimation, we have more data for the blog equation (Eq. (2)) than data for the outcome equation (Eq. (1)), as the latter data series begins only after launch. Our specified error structure is that all errors are iid, and the error distribution for \( \epsilon_t \) is invariant to launch (i.e., the model has an error structure that is univariate until launch and bivariate normal following launch). The intuition behind the identification is that some parameters in the model (e.g., \( \rho \)) are identified based on the likelihood that is generated from the bivariate normal error structure, whereas other parameters (e.g., \( \beta_1 \)) are identified based on the likelihood that is generated from both error structures.\(^{11}\)

4. Results

4.1. Parameter estimates

We initially focus on the market outcome equation (Table 9). Although we find that the coefficient of the total number of current blogs is not significant, we also find that the coefficient of the volume of current blogs that are classified as having a positive valence is positive and significant.\(^{12}\) This result is consistent with some previous findings for online reviews (e.g., in Dellarocas et al., 2007) but is inconsistent with other findings (e.g., Liu, 2006 and Duan, Gu, & Whinston, 2008) that find the number of positive reviews was not predictive of movie sales. Moreover, we find that the Blogstock variable is positive and significant. This result suggests that the volume of both current and past blogs is a predictor of market outcomes. In general, we expect that these effects are obtained because blog posts represent information and opinions regarding service experiences and product attributes by other consumers; thus, these posts are a rich source of insights to consumers in general.

The coefficient for GRP Adstock is positive and significant, and this finding is consistent with previous research for this category (e.g., Elberse & Anand, 2007). Next, we find that the coefficient for the interaction term between Blogstock and Adstock is positive and significant. This finding suggests that the two media (new and traditional) act synergistically; this result represents an interesting and hitherto undocumented finding. As expected, the lagged outcome terms (i.e., both the one-period lag and the stock variable) are positive and significant. With regard to the four words that were retained following the text-mining analysis, we find that the coefficient of “Viewed” is positive and significant, whereas the coefficient for “Award” is...

\(^9\) We also confirmed with our data provider(s) that the firms expected their advertising to have a direct effect on sales alone. In other words, the firms did not expect their advertising to influence the blogosphere.

\(^{10}\) The exact value of the carryover constant is estimated via a grid search with the best fit as the objective function. Details regarding this analysis are available from the authors upon request.

\(^{11}\) For computational ease, we use FGLS to estimate the model (Wooldridge, 2001, Chapter 7).

\(^{12}\) As a robustness check, we included only the number of blogs (i.e., we eliminated the number of positive and negative blogs from the specification). In this instance, the number of blogs became positive and significant.
negative and significant. This result provides additional process evidence that blogs affect market outcomes because blogs represent a direct and valid source of information regarding consumer opinion for other consumers.

With regard to the equation with the number of blogs as the dependent variable, we examine the pattern of results (Table 10) by observing the pre- and post-launch estimates. We first examine the effect of advertising on blog production. We find that advertising in the pre-launch period leads to a positive and significant increase in the number of blogs. Interestingly, this pattern changes in the post-launch period: advertising does not affect blog production. As an increasing number of consumers are able to sample a product (post-launch), bloggers are likely to be less affected by traditional advertising because they are able to rely on their own experiences and/or information regarding product attributes and performance. The results pertaining to the frequency of words also reinforce this finding: mentions of “Advertising” are positive and significant pre-launch (albeit at the p = 0.10 significance level) but not post-launch. The same pattern is observed for “Award”; thus, this result suggests that bloggers focus on different sets of information pre-versus post-launch. Finally, the word “Interesting” is also significant only during the pre-launch period. An interesting implication of these results is that managers could use the volume and content of blogs as leading indicators of the effectiveness of their advertising. In other words, if managers employ an ad campaign and find no effect on the blogosphere, this result could suggest that the campaign was not satisfying its desired objective. Again, these findings are novel and provide insight into the relationship between traditional and new media from both a raw count measure and a content measure.

We also find that the stock of blogs is predictive (positive and significant) of the number of blogs in a given period during both the pre- and post-launch periods. This finding also seems reasonable because anecdotal evidence suggests that the blogosphere “feeds off itself.” With regard to the effects of marketing outcomes on blogging, we find that there is generally some type of relationship between them. Specifically, a current audience size (sales) has a positive and significant effect on the number of blogs, whereas a cumulative audience has a negative and significant effect on the number of blogs. This trend may arise because movie audiences decline rapidly (e.g., as shown in Fig. 2), but the discussion of movies persists in the blogosphere. Moreover, a movie is typically a one-time consumption event and thus has a negative and significant effect on the number of blogs. The results pertaining to the frequency of words also reinforce this finding: mentions of “Advertising” are positive and significant pre-launch (albeit at the p = 0.10 significance level) but not post-launch. The same pattern is observed for “Award”; thus, this result suggests that bloggers focus on different sets of information pre-versus post-launch. Finally, the word “Interesting” is also significant only during the pre-launch period. An interesting implication of these results is that managers could use the volume and content of blogs as leading indicators of the effectiveness of their advertising. In other words, if managers employ an ad campaign and find no effect on the blogosphere, this result could suggest that the campaign was not satisfying its desired objective. Again, these findings are novel and provide insight into the relationship between traditional and new media from both a raw count measure and a content measure.

4.3. Managerial implications

Thus far, we have discussed our findings purely from a statistical perspective. However, it may be useful to translate these findings in a manner that quantifies the effect sizes from a managerial perspective. Therefore, we performed two simulations: the first simulation was intended to determine how managers could change resource allocation, and the second simulation was designed to examine how managers could use blog data to improve sales forecasts.

In the first simulation, we assumed that there were only three periods: two pre-launch periods and one post-launch period. Recall that blogging is outside of the control of managers. Therefore, we used a marketing instrument under managerial control in our data set—traditional TV advertising.13 In the experiment, we increased the Adstock variable by 1% in the first pre-launch period. The output that we measured was the percentage of increase in the size of the daily volume sold during the post-launch period.

As shown in Fig. 3, a one-percent increase in Adstock resulted in a 0.16% increase in the number of blogs during the following time period (the second pre-launch period). As a result of this increase in the Adstock variable, we found that the net increase in sales was 0.095%. An analysis of this overall increase due to traditional media versus new media suggested that the increase in Adstock caused an increase in the audience size of 0.022%, whereas the increased blogging activity caused an increase in the audience size of 0.073%. In other words, approximately three-fourths of the total short-term effect is attributed to blogs (the indirect effect). In addition to the short-term effect, we computed the long-term effect via a similar simulation by holding everything else constant. Fig. 4 illustrates the effect size over time. As shown in the figure, the effect decreases to approximately 10% of the short-term effect in twenty time periods (days) and trends to zero in approximately thirty time periods (days). Due to carryover, approximately 90% of the total long-term effect is derived from the indirect channel.

It is unclear why the indirect effect accounts for such a large proportion of the overall effect and whether this finding is an anomaly.

### Note

13. Note that our analysis is essentially correlational because it is difficult to prove causality based on our data. Therefore, our simulation must be considered an attempt to document the economic effects of firm actions rather than as a prescriptive guideline for managerial action.
this lack of clarity could arise partly because models that include both new and traditional media and their interaction have not been estimated previously. In addition, the audiences for blogs and TV advertising could overlap significantly for the movie category in Japan. Therefore, the context in which this result should be interpreted should become clearer as an increasing amount of research on this topic becomes available.

In the second simulation, we held out the last observation from each brand and re-estimated the model. We then used the model estimates for prediction and computed the difference in the predicted value and the actual data across all of the held-out observations. We used this method for the full model and for a restricted version of the full model in which the response coefficients for the number of blogs and the Blogstock variable were set to zero. Thus, the difference in predictions (based on the Root Mean Square Deviation metric or RMSD) between these two models shows the extent to which the use of blog data can improve sales forecasts. The full model shows a large increase in prediction ability as the RMSD decreases from 8.42 (for the restricted model) to 2.67.

These two simulations suggest that managers could reallocate marketing resources from traditional media if they find a significant indirect effect on market outcomes (i.e., via new media), as we have found. In other words, free media acts as a proxy for paid media. As demonstrated by previous research on resource allocation when there is a positive interaction between two forms of paid media (e.g., Naik, Raman, & Winer, 2005), this exercise may be valuable for managers. In addition, the incorporation of new media activity into sales response models is promising in terms of improvements in sales predictions.

Table 13

<table>
<thead>
<tr>
<th>Period</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All periods</td>
<td>5804.03</td>
<td>4288.31</td>
</tr>
<tr>
<td>Pre-launch</td>
<td>5823.75</td>
<td>4623.10</td>
</tr>
<tr>
<td>Post-launch</td>
<td>5791.48</td>
<td>4105.04</td>
</tr>
</tbody>
</table>

5. Cellular phone category

As mentioned previously, we were also able to obtain data for the cellular phone category that were similar to that for the movie product category. We obtained the monthly number of subscribers from the Japanese Telecommunication Carriers Association website. We used data on monthly subscribers from five Japanese cell phone companies for the eighteen-month period from November 2006 to April 2008. These companies were NTT DoCoMo with 53 million subscribers in April 2008, Softbank Mobile with 19 million subscribers, Au with 30 million subscribers, Willcom with 4.6 million subscribers, and EMOBILE (the only launch that we observed in this category in May 2007) with 0.5 million subscribers. Therefore, we use that date (i.e., May 2007) to define the pre- and post-launch periods. In our data set, the average number of new subscribers each month was 110,145, with a standard deviation of 120,692. Tables 13, 14 and 15 present the TV GRPs, the volume of blogs and the valence of blog data, respectively, for this category.

We followed the same text-mining procedure as described previously. However, as the number of posts related to cellular phone service was much larger than that for movies, we drew a 5% (260,000 posts) random sample for our analysis. Using the same criteria as for movies, we chose “Advertising,” “Interest,” “Subscribed,” and “Price” because these words (especially the latter three words) can be viewed as diagnostic of the service. Table 16 shows the proportion of blogs that use each of the chosen words. As expected, “Advertising” has a higher proportion of pre-launch mentions, whereas “Price” and “Subscribed” have a higher proportion of post-launch mentions.

The results for the model are presented in Tables 17 (for the marketing outcome equation) and 18 (for the blog equation). The relevant results (i.e., those related to new media) are generally consistent with those for the movie category. In the market outcome equation, the coefficient for the current number of blogs remains insignificant, whereas the coefficients for the number of current positive blogs and the Blogstock variable are positive and significant. As expected, the coefficients for “Subscribed” are positive, and those for “Price” are negative. Most importantly, we again find evidence of a synergistic relationship between traditional and new media (the interaction term is positive and significant).15 We also replicate our key findings for the blog equation. Specifically, we find that Adstock has a positive and significant pre-launch effect on blog production but has no post-launch effect. Moreover, in terms of process measures, the mention of “Advertising” is positive and significant during the pre-launch period but not during the post-launch period.

In terms of the model fit (detailed results are available from the authors upon request), we find that the inclusion of the blog variables

---

14 Note that EMOBILE launched a series of calling plans but did not change the prices and attributes of these plans for the duration of our data.

15 One contrasting finding for cellular phone service is that the coefficient for Adstock is negative and significant. Although this result may seem counterintuitive, a closer examination of the data patterns suggests the reason for this finding. Recall that, for this product category, we observe only one product entry (EMOBILE) out of five brands. The data suggest that, to defend against the new launch, the incumbents raised their total advertising spending by approximately 30% and 20% in the first and second periods following launch, respectively. However, sales remained virtually flat after launch (the new brand gained an insignificant amount of market share). This outcome, coupled with the use of the Adstock formulation, results in a situation of flat sales with increased advertising, leading to a negative coefficient of Adstock.
improves the fit both in and out of the sample to an even greater degree in the movie category. The simulation results are also similar.

Overall, the results from the cellular phone service category replicate the relevant results from the movie category; this outcome gives us confidence that our results are not idiosyncratic to the movie category and may thus be generalizable (across many product categories).16

6. Conclusion, limitations and directions for future research

This paper contributes to the limited but rapidly growing field of research on the information content and predictive power of new media in the presence of traditional media, especially in the context of new product launches. Using a unique data set from two product markets in Japan (a major new media market), we are able to combine data on market outcomes, traditional media (TV advertising) and new media (volume, valence and content of blogs) into a single source. Given the newness of this field, we adopt a primarily empirical approach in our analysis. We motivate the use of our specification by leveraging theoretical and empirical findings from previous research, and we hope that our work will lead to the development of a unifying theoretical framework for understanding user-generated content in future research.

We used a simultaneous equation model to capture the effect of new media on market outcomes and the effect of market outcomes on new media in the presence of traditional media. The first of our two key findings is that new media and traditional media act synergistically vis-à-vis market outcomes. To the best of our knowledge, this research is the first study to demonstrate this relationship. Our second key finding suggests that this relationship is much stronger during the pre-launch period than during the post-launch period, when consumers can rely more heavily on actual product performance. Finally, our paper represents one of the early attempts to use the content that is embedded in new media to provide "process" insight regarding our econometric findings via a systematic text-mining analysis.

Specifically, using data from the movie market in Japan, we find patterns showing clear relationships among traditional media, new media and market outcomes. In general, we find that the stock of past blogs is predictive of market outcomes, blogs and TV advertising, which act synergistically; that pre-launch advertising spurs blogging activity (which is predictive of marketing activity) but becomes less effective in inducing blogging activity post-launch; and that market outcomes also have affect blogging activity. Our text-mining results provide additional support for some of these findings. We are also able to replicate these findings with data from a different product category (i.e., cellular phone service); thus, this result suggests the potential generalizability of our findings. From a managerial perspective, we find that a 1% increase in the traditional marketing instrument (TV advertising) leads to an approximately 0.1% increase in market outcomes in the short term, with the majority of the increase being attributed to the increase in blogging activity that is generated by pre-launch advertising. By performing a second simulation, we also show that the use of blog data can lead to better sales forecasts.

Although our analyses are descriptive and primarily leverage correlational patterns, our results (and the second simulation above) suggest that blogs can be good predictors of market outcomes and that managers should include them in sales forecasting models. Although this implication is "passive", our results also suggest a more "active" implication. Essentially, by understanding the specific relationship between traditional and new media, managers can more adequately allocate resources to traditional media because of their ability to exploit the "multiplier" effect of traditional media on new media (as we showed in the first simulation above). Another useful implication that emerges from our findings is that blogging activity can be a surrogate measure of advertising effectiveness, which is typically difficult to gauge directly using traditional methods. Finally, managers may also find that rigorous text-mining analyses can provide insight on consumer evaluation and adoption of new products.

Our analyses also have some limitations (which are primarily driven by the nature of the data). First, as noted previously, the aggregate nature of our data increases the difficulty of offering micro-level causal explanations of our results. Although our text-mining analyses provide some insight into our findings, it would be beneficial to obtain datasets that link individual activity to market outcomes for a larger variety of new media. Second, our measures of new media are currently limited to blog post volume, content and valence. We do not possess any knowledge of the identity and demographics of the bloggers and the extent of linkage to and from a given blog (e.g., trackbacks). Third, it would be useful to extend our analysis to additional categories from Japan and other markets. This extension would allow for a detailed analysis of both types of results, including those that are systematically identical across categories and those that differ. Finally, our data do not contain information on all marketing instruments; thus, we use proxies (such as lagged sales in the case of distribution). We hope that future research will benefit from enhanced data to address these limitations.

---

16 We were also able to obtain data for the green tea drink product category. The data included five product launches. However, we were unable to obtain data pertaining to the content of blogs; thus, we could not conduct a text-mining analysis. However, we were able to replicate most of our results using a model that did not include any covariates based on text mining. Details are available from the authors upon request.
Stock variables are defined as in Eq.(4).

Two-sided significance test symbols: **p<0.001, *p<0.01, 'p<0.05, and 'p<0.10.

#Blog Positive and Negative are relative to #Blog neutral.

Stock variables are defined as in Eq. (4).

Acknowledgments

The authors would like to thank the Editor, the Area Editor and two anonymous referees for their support and encouragement; Lada Adamic; Eugene Anderson; Anocha Aribarg; Rajeev Batra; Gerry Davis; Fred Feinberg; Srinisvaragghan Srinaram; seminar participants at Temple University, Rutgers University, the University of Delaware and Erasmus University; and the participants at the 2010 Marketing Dynamics Conference in Istanbul, the 2009 YCCI Conference at Yale SOM, the 2008 Marketing Science Conference in Vancouver and at the Michigan Marketing PhD Camp 2008 for valuable feedback. The authors would also like to thank Buzz Research Inc. and Dentsu Inc. for providing the data. Onishi would also like to thank the Ross School of Business for its research support via the Spivey/Hall Fellowship. All correspondence may be addressed to the authors at hohnishi@kde.biglobe.ne.jp or pmanchan@umich.edu. Correspondence by regular mail should be addressed to the second author at the Ross School of Business, University of Michigan, 701 Tappan St., Ann Arbor, MI, 48109, USA. The standard disclaimer applies.

Appendix A. Robustness check with a three-equation system

To allow for a complete system, as mentioned in the text, we also estimated a simultaneous equation model in which we added a third equation to the system that is defined by Eqs. (1) and (2). This equation is given below:

\[
\text{ln} (TV GRPs) = \gamma_0 + \gamma_1 \text{Trend}_t + \gamma_2 \text{Season}_t + \gamma_3 \text{Brand}_t + \gamma_4 \text{#Customer}_t + \gamma_5 \text{ln} (\text{salestock}_t) - 1 + \text{Day of the week}_t + \epsilon_{t}.
\]

We found that the results from the sales volume equations and blog equations were the same between the two-equation models and the three-equation system. The advertising equation results (Table A5-a, b, and c) showed that the current TV GRPs were significantly correlated with the lagged TV GRPs. The results also suggest that Blogstock and Salestock were related to TV GRPs, especially in the post-launch period.

Table 18
Blog equation (cellular phone service).

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-19.61</td>
<td>7.68</td>
</tr>
<tr>
<td>Trend</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Brand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand1</td>
<td>1.82</td>
<td>0.34</td>
</tr>
<tr>
<td>Brand2</td>
<td>2.14</td>
<td>0.13</td>
</tr>
<tr>
<td>Brand3</td>
<td>1.95</td>
<td>0.13</td>
</tr>
<tr>
<td>Brand4</td>
<td>-0.65</td>
<td>0.27</td>
</tr>
<tr>
<td>Pre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#Blog(t−1)</td>
<td>0.09</td>
<td>0.19</td>
</tr>
<tr>
<td>#Blog Positive(t−1)</td>
<td>0.20</td>
<td>0.03</td>
</tr>
<tr>
<td>#Blog Negative(t−1)</td>
<td>-0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>#Blog Price(t−1)</td>
<td>0.44</td>
<td>0.22</td>
</tr>
<tr>
<td>#Blog Advertising(t−1)</td>
<td>0.15</td>
<td>0.20</td>
</tr>
<tr>
<td>#Blog Subscribed(t−1)</td>
<td>0.48</td>
<td>0.11</td>
</tr>
<tr>
<td>Blogstock</td>
<td>0.24</td>
<td>0.08</td>
</tr>
<tr>
<td>Adstock</td>
<td>0.20</td>
<td>0.03</td>
</tr>
<tr>
<td>Blogstock-Adstock</td>
<td>-0.03</td>
<td>1.1E-04</td>
</tr>
<tr>
<td>Post</td>
<td></td>
<td></td>
</tr>
<tr>
<td>#Blog(t−1)</td>
<td>2.80</td>
<td>1.26</td>
</tr>
<tr>
<td>#Blog Positive(t−1)</td>
<td>1.76</td>
<td>0.32</td>
</tr>
<tr>
<td>#Blog Negative(t−1)</td>
<td>-0.58</td>
<td>0.17</td>
</tr>
</tbody>
</table>
| #Blog Subscribed(t−1) | 0.71 | 0.42 | 1.69 |^
| #Blog Price(t−1) | 0.54 | 0.28 | 1.94 ^ |
| #Blog Advertising(t−1) | 0.12 | 0.20 | 0.63 |
| #Customer(t) | -0.06 | 0.11 | -0.59 |
| Blogstock | 0.10 | 0.18 | 0.57 |
| Salestock | 0.45 | 0.20 | 2.32 * |
| Adstock | -0.09 | 0.06 | -1.46 |
| Blogstock-Adstock | 0.02 | 6.3E-04 | 31.86 *** |
| Other product blogs | 1.84 | 2.1E-03 | 891.40 *** |

N=90.

Two-sided significance test symbols: **p<0.001, *p<0.01, 'p<0.05, and 'p<0.10.

#Blog Positive and Negative are relative to #Blog neutral.

Stock variables are defined as in Eq. (4).

Table A-1
Sales volume (audience) equation (movies).

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.06</td>
<td>0.61</td>
</tr>
<tr>
<td>Day of the week</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tue</td>
<td>0.40</td>
<td>0.07</td>
</tr>
<tr>
<td>Wed</td>
<td>0.93</td>
<td>0.05</td>
</tr>
<tr>
<td>Thu</td>
<td>0.38</td>
<td>0.05</td>
</tr>
<tr>
<td>Fri</td>
<td>0.49</td>
<td>0.05</td>
</tr>
<tr>
<td>Sat</td>
<td>1.07</td>
<td>0.05</td>
</tr>
<tr>
<td>Sun</td>
<td>0.98</td>
<td>0.05</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.92</td>
<td>0.10</td>
</tr>
<tr>
<td>Trend</td>
<td>-0.01</td>
<td>2.5E-03</td>
</tr>
<tr>
<td>Brand</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Movie1</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>Movie2</td>
<td>0.39</td>
<td>0.10</td>
</tr>
<tr>
<td>Movie3</td>
<td>0.42</td>
<td>0.07</td>
</tr>
<tr>
<td>Movie4</td>
<td>-0.07</td>
<td>0.08</td>
</tr>
<tr>
<td>Movie5</td>
<td>0.45</td>
<td>0.11</td>
</tr>
<tr>
<td>Movie6</td>
<td>3.0E-03</td>
<td>0.07</td>
</tr>
<tr>
<td>Movie7</td>
<td>0.14</td>
<td>0.06</td>
</tr>
<tr>
<td>Movie8</td>
<td>0.28</td>
<td>0.06</td>
</tr>
<tr>
<td>Movie9</td>
<td>-0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>Movie10</td>
<td>0.31</td>
<td>0.05</td>
</tr>
<tr>
<td>Movie11</td>
<td>-0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>#Audience(t−1)</td>
<td>0.28</td>
<td>0.11</td>
</tr>
<tr>
<td>#Blog(t)</td>
<td>-0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>#PositiveBlog(t)</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>#NegativeBlog(t)</td>
<td>1.0E-03</td>
<td>0.23</td>
</tr>
<tr>
<td>#Blog Viewed(t)</td>
<td>0.30</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table A-1 (continued)

<table>
<thead>
<tr>
<th></th>
<th>Full model (Proposed)</th>
<th>Three-equation model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>Std. Error</td>
</tr>
<tr>
<td>#Blog Positive(t)</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>#Blog Negative(t)</td>
<td>0.35</td>
<td>0.02</td>
</tr>
<tr>
<td>Advertising(t)</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>#Consumer(t)</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Stock variables</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

N = 1729.
Two-sided significance test symbols: ***p < 0.001, **p < 0.01, *p < 0.05, and p < 0.10.
#Blog Positive and Negative are relative to #Blog neutral.
Stock variables are defined as in Eq. (4).

Table A-2
Blog equation (movies).

<table>
<thead>
<tr>
<th></th>
<th>Full model (Proposed)</th>
<th>Three-equation model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>Tue</td>
<td>0.18</td>
<td>0.10</td>
</tr>
<tr>
<td>Wed</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>Thu</td>
<td>0.14</td>
<td>0.10</td>
</tr>
<tr>
<td>Fri</td>
<td>0.24</td>
<td>0.10</td>
</tr>
<tr>
<td>Sat</td>
<td>0.18</td>
<td>0.10</td>
</tr>
<tr>
<td>Sun</td>
<td>0.23</td>
<td>0.10</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.23</td>
<td>0.10</td>
</tr>
<tr>
<td>Trend</td>
<td>0.23</td>
<td>0.10</td>
</tr>
<tr>
<td>Movie1</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td>Movie2</td>
<td>0.19</td>
<td>0.04</td>
</tr>
<tr>
<td>Movie3</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie4</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie5</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie6</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie7</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie8</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie9</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie10</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Movie11</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Pre</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog Positive(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog Negative(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>Advertising(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Consumer(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog Positive(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog Negative(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
<tr>
<td>#Blog Viewed(t−1)</td>
<td>0.07</td>
<td>0.02</td>
</tr>
</tbody>
</table>

N = 1729.
Two-sided significance test symbols: ***p < 0.001, **p < 0.01, *p < 0.05, and p < 0.10.
#Blog Positive and Negative are relative to #Blog neutral.
Stock variables are defined as in Eq. (4).

Table A-3
Advertising equation (movies).

<table>
<thead>
<tr>
<th></th>
<th>Full model (Proposed)</th>
<th>Three-equation model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est.</td>
<td>Std. Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Day of the week</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Week</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Month</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Holiday</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Trend</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Brand</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>TV GRP(-1)</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Post</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>#Blog</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>#Consumer</td>
<td>0.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Stock variables</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

N = 1729.
Two-sided significance test symbols: ***p < 0.001, **p < 0.01, *p < 0.05, and p < 0.10.
#Blog Positive and Negative are relative to #Blog neutral.
Stock variables are defined as in Eq. (4).

References
The performance of global brands in the 2008 financial crisis: A test of two brand value measures

Johny K. Johansson a, Claudiu V. Dimofte b,⁎, Sanal K. Mazvanchery c

a McDonough School of Business, Georgetown University, Washington DC 20057, United States
b San Diego State University, San Diego CA 92182, United States
c Kogod School of Business, American University, Washington DC 20016, United States

Abstract

Previous literature has argued that high brand equity helps stabilize financial returns and reduce share price volatility. This research investigates how some of the strongest brands in the U.S. market fared in terms of financial performance during the Fall 2008 stock market downturn. Initial results using a financially based measure of brand value (Interbrand) show that, counter to expectations, these top brands did not outperform the market as a whole. However, the findings are in the hypothesized direction when an alternative, consumer-based brand equity measure (EquiTrend) is used to replicate the analysis. After first employing the three Fama–French factors to evaluate stock performance, we assess the added brand equity effect using both aforementioned measures. The consumer-based measure shows a significant incremental effect on stock performance after controlling for risk and financial fundamentals. Furthermore, this positive effect also applies to share volatility and firm betas. None of these effects hold for the financially based measure.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Relating marketing indicators to financial indicators of stock market performance and shareholder value has been the focus of several recent publications in the academic marketing literature (e.g., Mizik & Jacobson, 2009b; O'Sullivan, Hutchinson, & O'Connell, 2009; Srinivasan & Hanssens, 2009). These analyses have shown that some of the firm's customer-level assets, such as customer satisfaction, customer equity, and brand equity, have a significant impact on financial performance (Fornell, Mithas, Morgeson, & Krishnan, 2006; Krasnikov, Mishra, & Orozco, 2009; Kumar & Shah, 2009). Specifically, the brand equity of a firm brand has been shown to have a significant positive impact on stock market performance. For example, Barth, Clement, Foster, and Kasznik (1998) used a sample of 1204 brand value estimates from 1991 to 1996 and found them to be positively related to stock prices and returns. Madden, Fehle, and Fournier (2006) juxtaposed a portfolio of 111 firms' brands from the Interbrand list of most valuable brands between 1994 and 2001 to a benchmark market portfolio and observed higher returns and lower risk for the Interbrand set. Finally, Rego, Billett, and Morgan (2009) used data from 252 EquiTrend listed firms between 2000 and 2006 to show that high brand equity reduces volatility and, thus, the risk associated with a brand's stock.

Despite the published research on the topic, there is no strong agreement in the marketing literature on how to define and measure brand equity. Two basic approaches can be distinguished. Brand equity can be measured either at the consumer level (e.g., Aaker, 1996) or at the financial markets level (Simon & Sullivan, 1993). Researchers following the consumer-level approach view brand image, consumer affinity, and customer loyalty as the main drivers of brand value (e.g., Aaker, 1996, Keller, 2007). Examples of commercially available, customer-based brand measures include Harris Interactive’s EquiTrend measure and Young & Rubicam’s Brand Asset Valuator. Alternatively, researchers employing financially based measures (e.g., Madden et al., 2006) focus on financial metrics, such as projected revenues and return on investment, to determine a brand's net present value. Commercially available monetized values of brands include the Interbrand Brand Value measure and Millward Brown's BrandZ. Whereas the consumer-based approach defines brand equity according to levels of consumer engagement, the financially based approach essentially translates intangible assets into financial figures by assessing a brand’s ability to generate future earnings.

It is not a priori obvious that the two approaches will concur in terms of what brands are more valuable. In fact, the two corresponding measures employed in this research (EquiTrend and Interbrand)
only correlate at \( r = .22 \) (ns) during 2008 \( (N = 50) \). Previous literature does not explain the reasons for such a low correlation. For example, brands ranked high according to consumer mindset measures supposedly command deep customer loyalty and are highly visible in consumer markets. These benefits ensure steady revenue levels and reduce firm idiosyncratic risk as well as its cost of capital (Rego et al., 2009). Alternatively, brands that fare well on financial markets-based measures tend to belong to firms with greater revenue streams and more predictable earnings (i.e., the so-called “large cap firms”). Large cap firms have a wider shareholder base and more prominent research coverage among analysts (Brennan & Subrahmanyam, 1996; Chordia, Subrahmanyam, & Anshuman, 2001). Thus, both types of measures suggest lowered risk and possibly higher returns for high equity brands. To date, no empirical comparison of the two types of measures has been undertaken. The purpose of this research is to present one such comparison.\(^1\)

The effect of brand equity on stock performance has been shown to be particularly beneficial in an economic downturn, when firms face reduced consumer demand (Deelersnyder, Dekimpe, Sarvary, & Parker, 2004) and thus earn lower revenues and profits (Srinivasan, Lilien, & Sridhar, 2011). In an economic crisis, firms with high brand equity would thus be more likely to sustain revenues than firms with lower equity. Furthermore, because investors are likely to search for less risky investments in such an environment, high equity brands should become particularly attractive as “safe harbor.” One would therefore expect the share prices of the strongest brands to lose less than those of weaker brands in an economic downturn. The financial crisis of 2008 and the stock market downturn in the Fall of 2008 (i.e., the September to December period when the market lost over 30% of its value) offer an ideal opportunity to test this general hypothesis.

In what follows, we analyze the stock market performance of 50 of the Interbrand Top 100 global brands from September to December of 2008. The 50 brands selected represent the subset where brand and stock market shares are directly related, and for which we have corresponding EquiTrend measures. We compare the Fall 2008 stock returns for the selected brands against the rest of their industries and the overall market. Initial results are consistent but not encouraging: shares of the highly rated brands did not drop significantly less than the market. The study then uses the three-factor Fama–French model (Fama & French, 1993) to account for differences in market-based risk for the 50 brands. The model shows that brands scoring the highest on brand equity as measured by EquiTrend do in fact show superior performance, in line with the basic hypothesis. We then introduce the financial fundamentals in order to avoid spurious inferences and to identify the incremental contribution from EquiTrend brand equity. For each of three dependent variables (stock returns, volatility, and beta) we investigate the additional explanatory power of brand value by introducing the Interbrand and the EquiTrend scores associated with each stock. Throughout the analyses we test for omitted variables and model misspecification by introducing additional variables and testing alternative functional forms (including an explicit consideration of the Fama–French market factors). Results are robust: within the sample, the EquiTrend measure is more successful than the Interbrand measure in identifying the brand equity that matters.

The rest of the article is organized as follows: the next sections develop the hypotheses and describe the data and the methodology used. A subsequent section presents the results and interpretation of the findings. The paper concludes with a discussion section, which includes an explanation of the differential impact of the two measures.

2. Hypothesis development

The year 2008 offers a dramatic example of a financial crisis. Over that year, the S&P 500 Index lost 38.5% and the Dow Jones Industrials (DJIA) average dropped 33.8%. The vast majority of stocks (almost 9 out of 10 of those in the broader S&P 1500 Index and more than 90% of those in the S&P 500) lost value during the year. On average, losing stocks dropped more than 40% of their value and almost $7 trillion in market value was wiped out. Shares of large firms with value-priced stocks, generally considered a safer part of the market, lost 38% of their value as measured by the Vanguard Value exchange traded fund. The 4-month period from early September through the end of December was particularly turbulent and included notable events such as the federal takeover of Fannie Mae and Freddie Mac, Lehman Brothers’ bankruptcy, and AIG’s $85 billion bailout. The Volatility Index (VIX) of the Chicago Board Options Exchange, known as the market’s “fear index,” reached an intraday all-time high of 89.53 during the year’s 4th quarter.

For the comparison of the two brand equity measures during the crisis, we focus on the critical period between September and December of 2008 (see Figure 1). If brand equity lowers risk, the effect should manifest itself most clearly in this time of great turbulence. The key comparison to make is which of the two measures better protected stock returns in this period.

2.1. Hypotheses

Brand equity is the added value that a brand name and its associated logo confer upon a product or service. For example, Aaker defines brand equity as “a set of brand assets and liabilities linked to a brand, its name and symbol, that add or subtract from the value provided by a product or service to a firm and/or that firm’s customers” (Aaker, 1991, p.15). The assets and liabilities “can be usefully grouped into five categories: brand loyalty, name awareness, perceived quality, brand associations in addition to perceived quality, and other proprietary brand assets—patents, trademarks, channel relationships, etc.” (Aaker, 1991, p.16).

A “strong” brand is one that can sustain and raise high positive brand equity over time, maintaining customer loyalty and successfully defending itself against competitive encroachment (Aaker, 1996). A brand that possesses high equity should be able to sustain both a price premium and a relatively stable revenue stream. The awareness associated with high equity will reduce consumer search costs and should facilitate repeat purchases (Kamakura & Russell, 1993). Further, the loyal consumer will be less susceptible to competitor appeals and will do less comparison shopping. In addition, the association of brand equity with high perceived quality will increase customer satisfaction and reduce the incentive to consider brand substitution (Berthon, Hulbert, & Pitt, 1999; Chaudhuri & Holbrook, 2001; Oliver, 1997).

Past research demonstrating the positive relationship between brand equity and share prices also suggests that investors recognize the positive effects (e.g., Aaker & Jacobson, 1994; O’Sullivan et al., 2009). This recognition is partly attributed to reputation, as investors tend to prefer well-known brands and brands with higher advertising spending (Joshi & Hanssens, 2010; McAlister, Srinivasan, & Kim, 2007). The share prices of high equity brands will rise to incorporate the advantages and the volatility of the shares should decrease to reflect the lower risk (Rego et al., 2009). Although the evidence on the extent to which the stock market prices efficiently incorporate high brand equity is still tentative (Mizik & Jacobson, 2008; O’Sullivan et al., 2009; Srinivasan, Pauwels, Silva-Risso, & Hanssens, 2009), the effect on the risk reduction is well established.
Research has shown that scoring highly on brand equity measures is associated with reduced systematic and unsystematic risk in the stock market (e.g., Rego et al., 2009; Srinivasan & Hanssens, 2009). Systematic risk refers to the variability in a firm’s stock returns due to the variation in the market as a whole. Unsystematic risk is the variability due to factors specific to the firm. Strong brand equity will help insulate the brand from market-level declines (Rego et al., 2009). One would therefore expect brands scoring highly on brand equity measures to show less sensitivity to market drops. This would mean that the market beta as computed during the crisis should be lower for high equity brands. Formally put:

**H1.** High equity brands will show lower betas than low equity brands in a downturn.

Brand assets are firm-specific, and thus their effect should be unique to the specific brand. This suggests that their effects are primarily idiosyncratic and distinct (see Aaker, 1996; Keller & Lehmann, 2006). Because brand equity values are sustained over time (Aaker & Jacobson, 1987), a reasonably efficient stock market would incorporate such advantages in share prices (see Rego et al., 2009). High brand equity simply means higher stock values. However, the stable earnings of stronger brands may make high equity brands particularly attractive during a severe downturn. High equity brands would then show more resistance to market-level shocks than other brands (Rego et al., 2009). This suggests that the volatility of brands scoring highly on brand equity measures during a crisis would be lower. Formally put:

**H2.** High equity brands will show lower volatility than low equity brands in a downturn.

The reduction in beta and volatility associated with high brand equity would then suggest that brands scoring highly on brand equity measures would show less of a drop in share values in a downturn. Formally put:

**H3.** The decline in stock returns will be lower for high equity brands than for low equity brands in a downturn.

### 3. Data

#### 3.1. Stock returns

The typical measure of a stock’s performance over a period is its return, usually calculated as the percent change of its share price over the period. We used COMPUSTAT daily closing prices on the NYSE, ASE, or NASDAQ for the period between September 1 and December 31, 2008. The return was then computed as the share price at the end of December 31, minus the price at the end of September 1, and divided by the price at the end of September 1. To obtain percentages, we multiplied by 100. Our basic expectation was that firms with strong brands (i.e., with high scores on brand equity measures) would show less of a drop in returns over the four-month period.

#### 3.2. Share volatility

The volatility of the stocks during the period was calculated using standard methods. Counting the trading days when the markets were open during the four months yielded a time series of 85 daily prices for each firm. From these prices, we computed 84 daily returns for each firm. The standard deviation of this series of returns was then used as a measure of share volatility. Our basic expectation was that shares of brands scoring highly on brand equity measures would show lower volatility.

#### 3.3. The betas

Finally, to calculate the firm’s reaction to the market drop during the period, we first computed the daily returns in the S&P 500 index for the period. Following the CAPM model, we then regressed daily returns against the market returns minus the risk-free returns to estimate the “beta” of the stock (Fama, 1970; Fama & French, 1993). Because we posit that high brand values insulate firms from a down market, we expected that the betas for firms with brands scoring highly on brand equity measures to be below 1.0 and the betas for firms with brands scoring lower on brand equity measures to be above 1.0, suggesting greater vulnerability to market swings.3

#### 3.4. Brand equity

The brands selected for the study were all part of the “100 Best Global Brands” ranking published by Interbrand4 in September

---

2 A common alternative measure involves taking the natural logarithm of price at the end of the period minus the natural logarithm of price at the beginning of the period. The two procedures yield very similar results.

3 Whether the beta and the volatility measures as calculated here are truly measures of “risk” is of course debatable. As risk measures, they should in principle be computed before the actual crisis occurred. They are useful here mainly as indicators of the degree to which the brand shares reacted to the overall market (the betas) and the degree to which brand equity might have made share prices relatively stable (volatility).

4 See [www.interbrand.com](http://www.interbrand.com) for the full list of the 100 brands. The rankings were based on brand data collected during the 12 months prior to June 30, 2008. How Interbrand scores are calculated is explained in Appendix A. Interbrand scores are valued in U.S. dollars and termed “brand values,” not strictly measures of “brand equity.”
Table 1a
Sample of 50 brands with brand scores.

<table>
<thead>
<tr>
<th>Brand</th>
<th>Interbrand value</th>
<th>EquiTrend value</th>
<th>Brand</th>
<th>Interbrand value</th>
<th>EquiTrend value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td>7.02</td>
<td>41.41</td>
<td>IBM</td>
<td>59.03</td>
<td>61.09</td>
</tr>
<tr>
<td>Amazon.com</td>
<td>6.43</td>
<td>67.84</td>
<td>ING</td>
<td>3.77</td>
<td>60.69</td>
</tr>
<tr>
<td>American Express</td>
<td>21.94</td>
<td>53.88</td>
<td>Intel</td>
<td>31.26</td>
<td>66.17</td>
</tr>
<tr>
<td>Apple</td>
<td>13.72</td>
<td>60.24</td>
<td>J.P. Morgan</td>
<td>10.77</td>
<td>53.13</td>
</tr>
<tr>
<td>Avon Products</td>
<td>5.26</td>
<td>51.48</td>
<td>Johnson &amp; Johnson</td>
<td>3.58</td>
<td>75.30</td>
</tr>
<tr>
<td>BP</td>
<td>3.91</td>
<td>57.86</td>
<td>Kellogg's</td>
<td>9.71</td>
<td>68.22</td>
</tr>
<tr>
<td>Canon</td>
<td>10.88</td>
<td>65.98</td>
<td>Marriott</td>
<td>3.50</td>
<td>62.59</td>
</tr>
<tr>
<td>Citibank</td>
<td>21.31</td>
<td>60.17</td>
<td>McDonald's</td>
<td>31.05</td>
<td>65.18</td>
</tr>
<tr>
<td>Coca-Cola</td>
<td>20.17</td>
<td>51.92</td>
<td>Microsoft</td>
<td>59.01</td>
<td>71.40</td>
</tr>
<tr>
<td>Colgate</td>
<td>6.44</td>
<td>67.58</td>
<td>Motorola</td>
<td>3.72</td>
<td>59.08</td>
</tr>
<tr>
<td>Daimler-Benz</td>
<td>25.58</td>
<td>55.75</td>
<td>Nike</td>
<td>12.67</td>
<td>62.90</td>
</tr>
<tr>
<td>Dell</td>
<td>11.70</td>
<td>62.66</td>
<td>Nokia</td>
<td>35.94</td>
<td>55.99</td>
</tr>
<tr>
<td>Disney</td>
<td>29.25</td>
<td>66.92</td>
<td>Oracle</td>
<td>13.83</td>
<td>50.29</td>
</tr>
<tr>
<td>eBay</td>
<td>7.99</td>
<td>63.39</td>
<td>Panasonic</td>
<td>4.28</td>
<td>62.29</td>
</tr>
<tr>
<td>FedEx</td>
<td>3.36</td>
<td>67.08</td>
<td>Pepsi</td>
<td>13.25</td>
<td>69.98</td>
</tr>
<tr>
<td>Ford</td>
<td>7.90</td>
<td>55.90</td>
<td>Philips</td>
<td>8.33</td>
<td>61.48</td>
</tr>
<tr>
<td>Gap</td>
<td>4.36</td>
<td>54.47</td>
<td>Research in Motion</td>
<td>4.80</td>
<td>54.07</td>
</tr>
<tr>
<td>GE</td>
<td>53.09</td>
<td>70.60</td>
<td>SAP</td>
<td>12.23</td>
<td>45.88</td>
</tr>
<tr>
<td>Google</td>
<td>25.59</td>
<td>73.53</td>
<td>Sony</td>
<td>13.58</td>
<td>67.62</td>
</tr>
<tr>
<td>Harley Davidson</td>
<td>6.65</td>
<td>79.27</td>
<td>Starbucks</td>
<td>3.88</td>
<td>57.64</td>
</tr>
<tr>
<td>Heineken</td>
<td>19.08</td>
<td>57.68</td>
<td>Toyota</td>
<td>34.05</td>
<td>63.52</td>
</tr>
<tr>
<td>Honda</td>
<td>23.51</td>
<td>64.89</td>
<td>UPS</td>
<td>12.62</td>
<td>72.83</td>
</tr>
<tr>
<td>HSBC</td>
<td>13.14</td>
<td>49.73</td>
<td>Yahoo!</td>
<td>5.50</td>
<td>71.92</td>
</tr>
</tbody>
</table>

* Brand values in billions of U.S. dollars (see Appendix A).
* Brand equity scores (see Appendix B).

For the portfolio consisting of these particular 50 stocks, the results were not in line with our hypothesis that these brands would beat the market. During the period between September 1 and December 31, 2008, the EW S&P 500 dropped 34.65%. The average drop during the same period for the 50 firms in our sample was higher at 35.62%. A t-test showed the difference to be non-significant ($p = .69$) but still surprising because it is in the “wrong” direction. We decided to eliminate the 8 brands representing institutions from the exposed financial sector, which produced slightly better results, with a drop of 32.28% for the remaining 42 brands; however, this result was not significantly better than the market ($p = .27$). In any case, this test is inappropriate because the S&P index of course includes the financial firms.\(^7\)

We then tested a weighted average of the 50 brands against the common weighted S&P 500 index. The weighted S&P 500 index dropped 30.27% during the four months, slightly less than the equal weights index. The weighted average for the portfolio of 50 brands with their relative market capitalization as the weights showed a mean drop of 32.92%, which was still in the “wrong” direction but not significant ($p = .64$). Eliminating the financial brands showed some improvement. The weighted average drop for the non-financial brands was 30.10%, approximately the same and slightly better than the market (though not significantly so). The damage incurred by the financial brands was clearly an important driver in the weak overall performance.\(^8\)

We then examined possible industry effects. The 50 brands may have come from particularly exposed industries, as was the case for the financial institutions. Industry effects in stock market analyses are often captured by using a dummy variable for each industry (e.g., Rego et al., 2009). In the present case, where the sample was limited in size and several industries were represented by very few brands, such a procedure was not feasible. Instead, using Google Finance, we grouped the 50 brands into 30 industries. The number of firms in an industry varied considerably, from as few as 4 (air couriers) to as many as 392 (regional banks). We then used the Wharton Research Data Service (WRDS), which draws on Compustat, to extract the share prices for each of the brands in the industry from September 1 to December 31, 2008. The percentage return was computed for each brand in the industry. A weighted average was then used to calculate the mean daily return for the industry in the period, the weights based on the “market cap” of the firm (its share price multiplied by the number of shares outstanding). The result was 30 time series of 84 observations each for the mean industry returns during the period. These data allowed us to estimate “industry betas” by regressing industry returns against the S&P 500. We also computed an adjusted set of industry mean returns, where in turn each of the 50 brands was excluded from the computation (to control for the fact that when one firm is a large player in an industry with few firms, its returns will dominate the industry average).

The mean drop for the 30 industries (excluding the focal brand) was 29.76%, close to the market using the weighted S&P 500 index. The industries were not unrepresentative or unique. Remarkably, brands scoring highly on Interbrand dropped more than their respective industries, where many weaker brands were presumably represented. The correlation between the returns of the 50 firms and the industry returns was positive but low ($r = .07$, $p = .63$).

\(^5\) The EquiTrend score calculation is described in detail in Appendix B.
\(^6\) For an evaluation of the differences between the two S&P indices, see Zeng et al. (2010). Weighted results were also computed in several instances, especially at the beginning of the analysis to determine whether the results changed much. The results were very similar and are available upon request. For the period under investigation, the two S&P indices correlated at $r = .99$.

\(^7\) There is a potential confound in the fact that the market index contains many of the Interbrand stocks. The S&P 500 includes 36 of the 50 brands in Table 1 (the foreign brands are excluded). We eliminated these 36 brands from the S&P 500 index to compute the index. The results reported here were virtually unchanged. They are available upon request.

\(^8\) A natural question here is whether the analysis should only cover the non-financial brands. The problem with such an approach is that the S&P 500 index does include the financial brands as well. Also, 13 of the top 100 Interbrand brands in 2008 were financial brands.
suggesting that the brands scored highly by Interbrand were not greatly influenced by what happened in their industries overall.

5. Analyzing risk and returns

5.1. Methodology

Although it was surprising to find the 50 brands performing worse than the market as a whole, the analysis needed to control for the particular risk factors of those brands. Although they all belonged to the Interbrand top 100, various risk factors may have made these stocks particularly vulnerable. To control for this, we applied the Fama–French three-factor model, which is a well-established method for conducting financial analysis of marketing effects (Fama & French, 1993; Srinivasan & Hanssens, 2009). The Fama–French three-factor model explains daily returns for a stock as a function of three risk-related factors: the overall market return adjusted for the risk-free (Treasury bill) return \((R_m - R_f)\), the difference in returns for small versus big stocks (SMB), and the difference in returns for high book-to-market versus low book-to-market stocks (HML). To calculate the risk characteristics of each stock, we used the period from May 1 to August 31, 2008, the four months immediately preceding the crisis (see Fig. 1). Data on the three daily factors during that period were downloaded from Kenneth French’s publicly available online database. The time series of 86 daily returns for each individual stock was then regressed on the three factors using the Newey-West heteroscedasticity and autocorrelation-consistent covariance matrix estimator (Newey & West, 1987) to calculate the standard errors and t-statistics. We also used 3 lags, following the Newey and West (1994) optimum lag method (see Appendix C for details, including estimation equations and the properties of the relevant error terms).

This approach generated the beta and the SMB and HML coefficients for each stock over the four-month period preceding the crisis. Analogous to the two-step Fama and MacBeth (1973) method, which has been widely used in the empirical analysis of the cross section of stock returns of financial panel data sets (Jagannathan & Wang, 1998; Shanken, 1992), these coefficients were then used as regressors to explain the cross-sectional differences in the subsequent 4 months between September 1 and December 31, 2008. In this period, all 50 stocks faced a similar market crisis. However, the reaction in their share prices should have differed depending on the estimated Fama–French coefficients, which capture the reaction of share prices to market swings. Thus, to the extent that the observed share drops simply reflected systematic market and risk factors, the share drops should not be attributed to idiosyncratic factors, including brand equity. To test whether brand equity still played a role, in a subsequent step we introduced the brand equity scores for each brand as an additional regressor. If the systematic risk factors were sufficient to explain the differences in share drops over the period, there would be no added role for brand equity.

In the cross-sectional analysis, we also introduced and tested other idiosyncratic variables that may explain observed differences in return drops and eliminate any brand equity effect. To this end, we largely followed Rego et al. (2009) in identifying several relevant “financial fundamentals.” These financial fundamentals are described next (see Table 1b for relevant summary statistics).

<table>
<thead>
<tr>
<th>Brand</th>
<th>Change%</th>
<th>Beta</th>
<th>Volatility</th>
<th>Diversif.</th>
<th>Leverage</th>
<th>Credit</th>
<th>Diversif.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIG</td>
<td>−92.85</td>
<td>2.30</td>
<td>15.78</td>
<td>1967</td>
<td>89.97</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>Amazon.com</td>
<td>−37.01</td>
<td>1.16</td>
<td>5.68</td>
<td>1994</td>
<td>81.54</td>
<td>6.25</td>
<td>1</td>
</tr>
<tr>
<td>American Express</td>
<td>−54.34</td>
<td>1.39</td>
<td>6.53</td>
<td>1850</td>
<td>92.64</td>
<td>8.25</td>
<td>1</td>
</tr>
<tr>
<td>Apple</td>
<td>−48.64</td>
<td>9.04</td>
<td>5.06</td>
<td>1976</td>
<td>42.67</td>
<td>8.00</td>
<td>1</td>
</tr>
<tr>
<td>Avon Products</td>
<td>−44.53</td>
<td>8.55</td>
<td>4.81</td>
<td>1886</td>
<td>86.88</td>
<td>8.00</td>
<td>1</td>
</tr>
<tr>
<td>BP</td>
<td>−13.46</td>
<td>1.12</td>
<td>5.23</td>
<td>1889</td>
<td>59.91</td>
<td>9.00</td>
<td>2</td>
</tr>
<tr>
<td>Canon</td>
<td>−27.70</td>
<td>1.06</td>
<td>5.29</td>
<td>1937</td>
<td>30.30</td>
<td>9.00</td>
<td>2</td>
</tr>
<tr>
<td>Cisco</td>
<td>−31.37</td>
<td>1.02</td>
<td>4.54</td>
<td>1943</td>
<td>41.43</td>
<td>8.25</td>
<td>2</td>
</tr>
<tr>
<td>Citibank</td>
<td>−64.89</td>
<td>1.86</td>
<td>11.59</td>
<td>1812</td>
<td>94.81</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>Coca-Cola</td>
<td>−13.06</td>
<td>0.68</td>
<td>3.40</td>
<td>1886</td>
<td>86.96</td>
<td>8.25</td>
<td>1</td>
</tr>
<tr>
<td>Coke</td>
<td>−10.66</td>
<td>0.64</td>
<td>3.23</td>
<td>1806</td>
<td>76.30</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>Daimler-Benz</td>
<td>−34.99</td>
<td>1.35</td>
<td>6.49</td>
<td>1886</td>
<td>71.70</td>
<td>7.75</td>
<td>2</td>
</tr>
<tr>
<td>Dell</td>
<td>−50.84</td>
<td>0.82</td>
<td>4.83</td>
<td>1984</td>
<td>86.11</td>
<td>7.75</td>
<td>2</td>
</tr>
<tr>
<td>Disney</td>
<td>−30.06</td>
<td>1.11</td>
<td>4.91</td>
<td>1923</td>
<td>47.40</td>
<td>8.00</td>
<td>1</td>
</tr>
<tr>
<td>eBay</td>
<td>−42.00</td>
<td>1.00</td>
<td>4.90</td>
<td>1995</td>
<td>23.83</td>
<td>7.75</td>
<td>2</td>
</tr>
<tr>
<td>FedEx</td>
<td>−24.24</td>
<td>0.83</td>
<td>4.31</td>
<td>1971</td>
<td>43.33</td>
<td>7.00</td>
<td>2</td>
</tr>
<tr>
<td>Ford</td>
<td>−49.22</td>
<td>1.31</td>
<td>9.25</td>
<td>1903</td>
<td>97.48</td>
<td>4.75</td>
<td>1</td>
</tr>
<tr>
<td>Gap</td>
<td>−32.00</td>
<td>0.98</td>
<td>5.28</td>
<td>1969</td>
<td>45.47</td>
<td>6.25</td>
<td>2</td>
</tr>
<tr>
<td>GE</td>
<td>−43.22</td>
<td>1.11</td>
<td>5.35</td>
<td>1892</td>
<td>84.46</td>
<td>10.00</td>
<td>2</td>
</tr>
<tr>
<td>Google</td>
<td>−33.87</td>
<td>0.89</td>
<td>4.46</td>
<td>1998</td>
<td>10.44</td>
<td>10.00</td>
<td>2</td>
</tr>
<tr>
<td>Harley Davidson</td>
<td>−58.44</td>
<td>0.64</td>
<td>6.19</td>
<td>1903</td>
<td>58.01</td>
<td>8.25</td>
<td>1</td>
</tr>
<tr>
<td>Heinz</td>
<td>−27.76</td>
<td>0.58</td>
<td>4.06</td>
<td>1869</td>
<td>81.51</td>
<td>7.00</td>
<td>1</td>
</tr>
<tr>
<td>Honda</td>
<td>−31.93</td>
<td>1.05</td>
<td>5.69</td>
<td>1948</td>
<td>62.85</td>
<td>8.25</td>
<td>2</td>
</tr>
<tr>
<td>HP</td>
<td>−21.11</td>
<td>0.83</td>
<td>4.26</td>
<td>1939</td>
<td>56.57</td>
<td>8.00</td>
<td>2</td>
</tr>
<tr>
<td>HSBC</td>
<td>−37.98</td>
<td>0.85</td>
<td>4.35</td>
<td>1865</td>
<td>94.25</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>IBM</td>
<td>−28.92</td>
<td>0.73</td>
<td>3.32</td>
<td>1911</td>
<td>76.36</td>
<td>8.25</td>
<td>1</td>
</tr>
<tr>
<td>ING</td>
<td>−65.13</td>
<td>1.52</td>
<td>8.95</td>
<td>1991</td>
<td>96.99</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>Intel</td>
<td>−35.08</td>
<td>0.98</td>
<td>4.61</td>
<td>1968</td>
<td>23.16</td>
<td>8.25</td>
<td>2</td>
</tr>
<tr>
<td>J. P. Morgan</td>
<td>−19.13</td>
<td>1.42</td>
<td>7.68</td>
<td>1823</td>
<td>92.11</td>
<td>8.75</td>
<td>1</td>
</tr>
<tr>
<td>Johnson &amp; Johnson</td>
<td>−16.59</td>
<td>0.63</td>
<td>2.99</td>
<td>1886</td>
<td>46.49</td>
<td>10.00</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1b

Descriptives for the 50 brands.4

4 Variables as described in the Methodology section.

5.2. Analyzing returns and risk

Age

In the cross-sectional analysis, we also introduced and tested other idiosyncratic variables that may explain observed differences in return drops and eliminate any brand equity effect. To this end, we largely followed Rego et al. (2009) in identifying several relevant “financial fundamentals.” These financial fundamentals are described next (see Table 1b for relevant summary statistics).

- **Age.** Older firms are more established and have already proven capable to withstand disruption across time. In a sense, they are “survivors” that will be more likely to attract investors in a downturn. Following Rego et al. (2009), this variable was coded as “1” for firms less than 25 years old, “2” for firms between 25 and 50 years old, and “3” for firms 50 years or older.10

- **Leverage.** Leverage was computed as the ratio of long-term debt plus current liabilities to total equity and referred to the degree to which borrowed funds were used to operate a business. The debt/equity ratio should have a negative effect on returns in a crisis, when high leverage likely exposes investors to greater equity risk (Ferreira & Laux, 2007).

- **Credit rating.** The firm’s credit rating is one signal of how risky the stock is. Stronger ratings should provide confidence to investors (both bond holders and equity holders) in an economic downturn.

10 Analyses employing the continuous age variable find identical results.
Table 2a
Three regression models of share returns\(^a\) (N=50).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Base model</th>
<th>Interbrand</th>
<th>EquiTrend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>−.601 (.000)</td>
<td>−.606 (.000)</td>
<td>−.495 (.000)</td>
</tr>
<tr>
<td>SMB coeff.</td>
<td>.221 (.061)</td>
<td>.208 (.076)</td>
<td>.142 (.233)</td>
</tr>
<tr>
<td>HML coeff.</td>
<td>−.243 (.039)</td>
<td>−.158 (.033)</td>
<td>−.130 (.281)</td>
</tr>
<tr>
<td>Brand score</td>
<td>.148 (.194)</td>
<td>.306 (.023)</td>
<td></td>
</tr>
<tr>
<td>Adj. (R^2)</td>
<td>.44</td>
<td>.44</td>
<td>.50</td>
</tr>
<tr>
<td>AIC</td>
<td>404.41</td>
<td>404.49</td>
<td>400.24</td>
</tr>
<tr>
<td>BIC</td>
<td>412.06</td>
<td>527.12</td>
<td>409.80</td>
</tr>
</tbody>
</table>

\(^a\) The Beta, SMB, and HML regressors are coefficient estimates from the four-month period May 1 to August 31, 2008, immediately preceding the Fall 2008 period analyzed here. AIC and BIC criteria both favor the EquiTrend model. In addition, F-tests comparing the performance of the brand equity models to the base model show that the addition of Interbrand did not improve the model fit (\(F(1, 45) = 2.55, p = .11\)), but the addition of EquiTrend did so significantly (\(F(1, 45) = 5.58, p < .02\)). In this research, we follow the convention of using two-tailed tests unless otherwise stated.

For our study, credit rating was coded using data from the Standard & Poor’s 2008 credit rating score on a 10-point scale ranging from “10” for an AAA rating to “1” for DDD.\(^{11}\)

• **Diversification.** Diversification captures the number of different industries in which the firm operates. The more businesses in which the firm operates, the lower is potentially its stock’s risk. Following Rego et al. (2009), this variable was coded “1” for a single industry firm and “2” for a firm with more diversification.

The estimation method closely drew on standard cautions in financial analysis. The wide-ranging size differences between the brands easily lead to unequal variances in the observations (heteroscedasticity) and suggested the use of generalized least squares analyses. The size differences also created a need to reduce the influence of extreme observations and outliers (see Barth et al., 1998). To address these problems, we employed the feasible generalized least squares (FGLS) regression approach using the STATA statistical package. FGLS uses the square of the residuals from the initial OLS results as the diagonal entries in the variance-covariance matrix. This matrix is then used for a generalized least squares estimation, each observation weighed in inverse proportion to the square of its residual in order to control for heteroscedasticity and the impact of outliers (Judge, Griffiths, Hill, Luethkepol, & Lee, 1985). The resulting coefficients are both efficient and consistent.

5.2. Stock returns and brand values

We first analyzed the role of high brand values in the return performance of the 50 stocks.

We began by using the three Fama–French terms (i.e., the beta and the SMB and HML coefficients for each stock over the four-month period preceding the crisis, as described above) to control for the influence of systematic risk factors (see Table 2a). Not surprisingly, these factors captured the returns performance of the 50 stocks associated with our brands well. To assess the possible incremental effect of brand equity beyond these factors, we introduced the two brand value measures in turn, as extra predictors of stock returns. Whereas the addition of Interbrand did not significantly affect these returns, EquiTrend performed much better. The introduction of EquiTrend scores significantly improved the adjusted R-square and produced the best model fit according to both the Akaike (AIC) and Bayesian (BIC) information criteria for model selection. Thus, we found that during the Fall 2008 crisis, firms with higher EquiTrend brand scores had higher returns (or less negative returns) than their systematic riskiness would have suggested. This is consistent with the notion that brand equity lowers unsystematic risk.

The EquiTrend scores did not change for any brand during the four-month period, but stocks with higher brand scores performed better than the stock’s systematic risk would have predicted. These results offer support for the “safe harbor” interpretation, the notion that investors move towards stocks with higher brand scores in a crisis. No such effect was observed for the high Interbrand stocks. However, the Interbrand scores for 2008 were made public on September 19, 2008, and it would therefore be possible that they contained unanticipated changes that affected the performance of the stocks. To assess this possibility, we calculated the change in Interbrand scores between 2007 and 2008 and re-ran the analysis for Interbrand changes as well as 2008 scores. The results again showed no significant impact from Interbrand scores. Given the slightly smaller sample size (N=45), the results were not conclusive. However, the direction of the results was consistent; Interbrand scores did not influence investors significantly (see Table 2b).

We tested the above results for spuriousness by assessing what other brand idiosyncratic factors may have played a role. Omitted variables could possibly account for the brand equity effect uncovered. Table 2c presents the results of three alternative models, which include industry performance as well as the firm-specific variables described in the methodology section.

As the results in Table 2c show, none of the financial fundamentals significantly explained share price performance over the period. Firm age is a marginally positive factor, suggesting that older firms may have better weathered the storm. There is clearly a possibility of multicollinearity obscuring any single variable’s impact, but that matters less because the emphasis here is upon the incremental impact of the brand value scores. It is clear that the improvement from Interbrand scores is small and non-significant.\(^{12}\) In contrast, the EquiTrend scores do suggest a significant positive effect of high brand equity. Once again, the introduction of EquiTrend scores significantly improves the adjusted R-square and produces the best model fit. Even with the financial fundamentals accounted for, high EquiTrend scores give the associated brands a significant boost.

We attempted other model specifications to further test for the possibility of omitted variable bias. The objective was to see whether the Interbrand scores would enter significantly and whether the EquiTrend effect could be eliminated. For example, we tested

\(^{11}\) We are aware of the role of imperfect credit ratings in the crisis (e.g., Lowenstein, 2010). Nevertheless, to be prudent from a statistical standpoint, the credit scores need to be accounted for. In fact, in the Fall of 2008, investors were apparently very much influenced by faulty credit ratings.

\(^{12}\) Note that because the included 50 brands all have relatively high scores on the Interbrand measure, this does not mean that Interbrand scores overall have no impact on share prices. This is not an analysis of the complete range of Interbrand scores, and thus is not an assessment of the scores’ validity.
whether the portion of revenues coming from the North American or other global markets could significantly affect firm performance. Neither of the two variables entered the models significantly. The international diversification of the revenue stream did not greatly matter, showing that the crisis was indeed a global one. We also tested whether financial firms receiving funds from the Trouble Asset Relief Program (TARP) in November 2008 showed a significantly different pattern, again finding no particular shifts in brand effects. The initial results remained robust.

5.3. Share volatility and brand values

The next step in the analysis was to assess the impact of brand value on stock volatility. The hypothesis here was that brands scoring high on brand equity measures would show less volatility. The volatility in the market, measured as the standard deviation of the S&P 500 for the daily returns across the four months from September 1 to December 31, 2008, was 4.07% for the weighted S&P 500 measure.

We then ran three regressions against the financial brands, however, was clearly reflected in the average betas. Selecting out the financial firms resulted in an average firm beta of .93, which was significantly lower than 1.0 (p < .03). The non-financial industries similarly lowered their average beta to 1.06, not significantly different from 1.0. The 8 financial brands had an average beta of 1.61, with an average industry beta of 1.34, which were both significantly higher than 1.0. In terms of the betas, the non-financial firms were significantly less sensitive to the market, while their industries were slightly more exposed.

We next ran the three beta regressions with financial fundamentals included. The results were in line with the previous analyses (see Table 4). The Interbrand coefficient estimate was very low, while the EquiTrend scores again showed strong significance, which helped to insulate the stocks from downward market swings.

6. Results summary

6.1. The top 50 brands

Contrary to expectations, the 50 brands selected from the Top 100 Global Brands did not perform better than the market in the Fall 2008 crisis. The returns of the 50 brands did not outperform the market but also significantly different from the market (p = .12). The higher volatility of the financial brands, however, was clearly reflected in the average betas. Selecting out the financial firms resulted in an average firm beta of .93, which was significantly lower than 1.0 (p < .03). The non-financial industries similarly lowered their average beta to 1.06, not significantly different from 1.0. The 8 financial brands had an average beta of 1.61, with an average industry beta of 1.34, which were both significantly higher than 1.0. In terms of the betas, the non-financial firms were significantly less sensitive to the market, while their industries were slightly more exposed.

We next ran the three beta regressions with financial fundamentals included. The results were in line with the previous analyses (see Table 4). The Interbrand coefficient estimate was very low, while the EquiTrend scores again showed strong significance, which helped to insulate the stocks from downward market swings.

5.4. Firm betas and brand values

We next examined the firms’ market betas and the related industry betas. We expected the firms scoring high on brand equity measures to show lower betas. We calculated the 4-month betas for the firms and the industries using the 84 daily returns in the period from September to December, 2008. The average beta for the 50 firms over the four months was β = 1.04, which was not significantly different from 1.0 (p = .49), suggesting that these firms tended to follow the market, on average.

Table 2c
Three extended regression models of share returns* (N = 50).

<table>
<thead>
<tr>
<th>Independent</th>
<th>Base model</th>
<th>Interbrand</th>
<th>EquiTrend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta</td>
<td>−504 (.001)</td>
<td>−504 (.001)</td>
<td>−401 (.007)</td>
</tr>
<tr>
<td>SMB coeff.</td>
<td>.347 (.015)</td>
<td>.330 (.030)</td>
<td>.288 (.061)</td>
</tr>
<tr>
<td>HML coeff.</td>
<td>−.266 (.071)</td>
<td>−.243 (.135)</td>
<td>−.153 (.311)</td>
</tr>
<tr>
<td>Industry change</td>
<td>.070 (.545)</td>
<td>.069 (.555)</td>
<td>.052 (.633)</td>
</tr>
<tr>
<td>Age</td>
<td>.247 (.086)</td>
<td>.234 (.127)</td>
<td>.250 (.067)</td>
</tr>
<tr>
<td>Leverage</td>
<td>−.092 (.530)</td>
<td>−.104 (.486)</td>
<td>−.102 (.459)</td>
</tr>
<tr>
<td>Credit rating</td>
<td>.152 (.245)</td>
<td>.128 (.380)</td>
<td>.118 (.335)</td>
</tr>
<tr>
<td>Diversification</td>
<td>−.104 (.383)</td>
<td>−.107 (.378)</td>
<td>−.097 (.388)</td>
</tr>
<tr>
<td>Brand score</td>
<td>.049 (.724)</td>
<td>.296 (.033)</td>
<td>.296 (.033)</td>
</tr>
<tr>
<td>Adj. R²</td>
<td>.44</td>
<td>.43</td>
<td>.59</td>
</tr>
<tr>
<td>AIC</td>
<td>410.16</td>
<td>411.88</td>
<td>406.36</td>
</tr>
<tr>
<td>BIC</td>
<td>427.37</td>
<td>431.00</td>
<td>425.48</td>
</tr>
</tbody>
</table>

* AIC and BIC criteria both favor the EquiTrend model. Furthermore, F-tests comparing the performance of the brand equity models to the base model show that the addition of Interbrand did not improve the model fit (F(1, 40) = .22, p = .64), but the addition of EquiTrend did so significantly (F(1, 40) = 4.79, p = .04).

5.4. Firm betas and brand values

We next examined the firms’ market betas and the related industry betas. We expected the firms scoring high on brand equity measures to show lower betas.

We calculated the 4-month betas for the firms and the industries using the 84 daily returns in the period from September to December, 2008. The average beta for the 50 firms over the four months was β = 1.04, which was not significantly different from 1.0 (p = .49), suggesting that these firms tended to follow the market, on average. The 30 industry-to-market betas averaged 1.10, which was not significantly different from the market (p = .12). The higher volatility of the financial brands, however, was clearly reflected in the average betas. Selecting out the financial firms resulted in an average firm beta of .93, which was significantly lower than 1.0 (p < .03). The non-financial industries similarly lowered their average beta to 1.06, not significantly different from 1.0. The 8 financial brands had an average beta of 1.61, with an average industry beta of 1.34, which were both significantly higher than 1.0. In terms of the betas, the non-financial firms were significantly less sensitive to the market, while their industries were slightly more exposed.

We next ran the three beta regressions with financial fundamentals included. The results were in line with the previous analyses (see Table 4). The Interbrand coefficient estimate was very low, while the EquiTrend scores again showed strong significance, which helped to insulate the stocks from downward market swings.

6. Results summary

6.1. The top 50 brands

Contrary to expectations, the 50 brands selected from the Top 100 Global Brands did not perform better than the market in the Fall 2008 crisis. The returns of the 50 brands did not outperform the market but instead were slightly worse than the market average (although not significantly so). Hypothesis 3 is not supported. The 50 brands also performed worse than their industry peers, and the results suggest that these leading brands did not necessarily move with their respective industries.

14 The industry data used in these beta regressions did not exclude the focal brands. We wanted the betas to reflect the full set of firms in each industry. To follow standard procedure, we also used the weighted S&P 500 measure as the market index. Betas are typically calculated for longer periods of time, such as a year, but we wanted to focus on the four months specifically. Still, the correlation between the Fall 2008 betas and the betas published for the whole 2008 year was a very high r = .99.
The results for volatility and betas were also inconsistent with previous research. The 50 brands showed higher volatility and higher betas than the market, which was not significant in either case but still demonstrated a trend toward higher riskiness. Excluding the financial brands from the set showed the expected pattern with both volatility and betas, which indicated lower sensitivity for non-financial brands than for brands from the financial sector. However, because the S&P 500 market indices also include the financial brands, the overall results reject both H1 and H2.

6.2. Interbrand

The results were inconsistent with the argument that stocks of high-value brands (according to Interbrand) provide a safe haven for investors during turbulent financial times. During the Fall 2008 stock market drop, brands with high Interbrand scores did not perform better than the market. If anything, they performed worse. They also underperformed relative to the industries they belong to, where less prominent global and local brands were included. Controlling for financial fundamentals and other related variables showed no significant impact on the 50 brands’ financial performance from brand values.

In terms of volatility, similarly surprising results were obtained. The 50 brands were more volatile than both the market and their industries, against expectations that high brand equity would reduce risk. As for betas, the 50 brands did not perform significantly worse than the market or their industries, but they did not do better either. Eliminating the financial firms lowered the betas significantly below 1.0. However, once financial fundamentals were controlled for, the Interbrand effect was not significant.

6.3. EquiTrend

The EquiTrend results generally show a much more positive picture of brand equity. Among the 50 brands, higher brand equity according to EquiTrend performed according to our initial hypotheses. Not only did high EquiTrend scores help lower volatility and the betas significantly, but the scores also showed a significantly positive impact on the share prices over the period. Brand equity as measured by EquiTrend helped lower riskiness and also helped limit overall losses in the Fall 2008 crisis.

It is important to note here that the 50 brands were selected from Interbrand’s Top 100 global list. If EquiTrend is superior to Interbrand, a selection of top EquiTrend brands would presumably outperform the market. To test this conjecture, we selected the top 50 EquiTrend brands from the industries previously analyzed. The simple average drop of these 50 brands between September and December of 2008 was 29.48%. This was significantly less than the EW S&P 500 drop of 34.65% (p<.02), which supports our present findings.

7. General discussion

This section will first address several issues of methodology and assess whether the observed results are valid and reliable. We will then discuss the possible explanations for the divergent results for the two brand equity measures.

7.1. Sample selection

The sample of 50 brands selected is small and of limited representativeness. One omission consists of the brands from conglomerate firms such as P&G and Unilever; another is the limit of shares listed on American stock exchanges; and a third is the exclusion of privately held firms. These brands are clearly not representative of the entire brand universe, whether in the U.S. or elsewhere. Nevertheless, the selected brands do include some of the most prominent global brands, and the aim was to evaluate how the best brands performed in the crisis. If these brands did not compare well to the vast majority of brands, it is difficult to believe that brand value is very important in stock market performance.

7.2. Measure selection

One measurement issue was the shift between equal weight and market cap weighted performance measures. We provided both in several instances. The main argument in favor of weighted measures is the use of market cap weights in the typical S&P comparisons. However, because equal weights have been shown to be preferable when markets are inefficient in pricing factors (e.g., Zeng, Dash, & Guarino, 2010,) and because they avoid introducing additional noise when weights fluctuate widely, equal weights may be preferable in this instance. In any case, as our findings show, the results are very similar.

7.3. Financial brands

The results clearly indicate that the Fall 2008 crisis hit financial firms the hardest. This is of course well known. The question is whether our results are mainly a result of including the financial institutions in our analysis. We think there is very little choice, given that our benchmark market indicators all include some or all of the financial firms. Nevertheless, we have shown that our results, even without the financial firms, fall into the same pattern. However, with a smaller sample size the statistical significance tends to be lower.

7.4. Time period

The chosen 4-month time period is limited. The cutoff dates of September 1 and December 31 2008 are necessarily arbitrary, although they did encompass the main portion of the crisis. The cutoff two weeks before the Lehman bankruptcy seemed logical because it was considered a significant event and because we wanted to include a period sufficiently long for the aftermath to play out. Once a longer period is addressed, many other factors need to be considered. However, we wanted to examine financial performance in a period where the market was clearly not in equilibrium and thus brand effects could enter prominently. Doing so, however, imposes a limit on the generalizability of the analysis; in equilibrium, the brand effect question revolves mainly around whether brands are properly valued by the market (see Mizik & Jacobson, 2009a).

7.5. Misspecification

The regression results could possibly be affected by misspecification, including omitted variables. Several additional financial and
related variables were tested. As noted earlier, foreign sales percentage was assessed, as was the use of TARP funds. We introduced alternative transformations of the financial fundamentals, including curvilinear relationships, by adding squared terms. We tested log transformations for the Interbrand independent variable and also used the Interbrand rankings instead of the dollar values. We also expanded on the brand scores by introducing “brand leadership” variables, drawing on the market share rank for the brands in their respective industries. Because no significant changes appeared in the brand equity effects, we retained the simple results presented in the tables.

7.6. Estimation method

The separate brand estimations during the calibration period (May 1–August 31, 2008) followed standard Newey–West methods for time series analysis. The feasible least squares method used for the cross-sectional analysis addresses some of the typical problems in cross-sectional stock data. It controls for heteroscedasticity and limits the influence of extreme observations and outliers (Judge et al., 1985). The introduction of industry averages helps control intra-industry error correlations, which are likely because different industry sectors are not necessarily affected to the same extent by a downturn. The results were stable across methods (OLS versus robust regression versus FGLS) and did not seem to be attributable to any method problems. In these data, the uncovered effects persist.

7.7. Brand equity

If the results hold, how can one explain why the EquiTrend brand equity scores possess more diagnostic information than the Interbrand brand value measure? Our first test was to check how the two measures correlated; as previously mentioned, the correlation was a meager $r = .22$ (ns). We next checked the degree to which the measures favored or discriminated against certain types of businesses. Brands are likely to be more important in some sectors than in others, for example, perhaps more important in the consumer goods sector than in the business-to-business sector (see Fischer, Voelckner, & Sattler, 2010). In particular, we asked whether the assessment of brand equity in financial institutions showed some systematic bias between the two measures. Financial institutions averaged a score of 53.27 in the EquiTrend measure, which was significantly lower ($p < .03$) than the average of 63.00 for the non-financial brands. However, the Interbrand measure showed a similar difference, from $17.5$ billion for non-financials to an $11.1$ billion average for financial brands, which was not significant at $p = .29$ but directionally consistent. We tried other splits (consumers versus durables, services versus products, B2B versus B2C) but the results were similar. The measures do not appear to differ in their assessment of brand equity by industry.

How strongly financial measures of brand equity should correlate with consumer-based measures is unclear. As we have observed, both measures have been used in past research to identify brand equity effects in the stock market, but there is no published empirical research directly comparing the two. Some previous research has put forward dimensions of brand value similar to the Interbrand–EquiTrend distinction. For example, Kamakura and Russell (1993) presented a dichotomy composed of Brand Value (quality perceptions discounted for price and advertising expenditure levels) and Brand Intangible Value (consumer brand name associations). Similarly, Francois and MacLachlan (1995) distinguished between measures of “brand equity,” which are largely financially oriented, and measures of “brand strength,” which tend to originate in consumers’ brand experiences and reactions. The value of financially based data as a measure of “brand equity” is questioned by Aaker (1996, p. 314), who notes that Interbrand scores are brand values, not equity measures. Overall, there seems to be agreement in the literature that financial and customer-based measures tap into different brand dimensions, whether they are called brand equity or brand value.

The low correlation found here clearly supports this assumption.15

8. Conclusion

The likely explanation for the differing performance of the two measures lies in the way equity is captured. The EquiTrend brand equity measure assesses consumer allegiance to a brand and is thus largely exogenous to the stock market. In contrast, Interbrand brand values partially rely on financial projections, which are necessarily predicated on specific assumptions about future growth. If the assumptions used to generate financially projected brand values no longer hold, the calculated brand values will not surprisingly be misleading. In contrast, the customer-based measures would point to a sustainable edge in the marketplace, even in, or perhaps particularly in, a down economy.

The authors acknowledge the generous assistance of Harris Interactive in providing access to their EquiTrend dataset and the financial support from Georgetown University’s Capital Markets Research Center. We are also thankful for the guidance provided by the editor, area editor, and three anonymous reviewers, as well as for helpful comments from Kimberly Cornaggia, Bill Droms, Brice Dupoyet, Natalie Mizik, and Keith Ord. Finally, we are particularly grateful to Brata Yudha for his assistance with the data collection.

Appendix A. Interbrand value calculations

The Interbrand formula deducts the following from total company earnings: (1) brand sales costs, (2) marketing costs, (3) overhead costs, including depreciation, (4) a charge for capital employed, and (5) taxes. The result of these deductions is then adjusted to account for the role of the brand in driving demand to determine what proportion of intangible earnings is attributable to the brand. The resulting brand earnings are then further adjusted by brand strength.

Brand strength involves several factors: (1) leadership (25%), or the brand’s ability to dominate a market (positive factor); (2) stability (15%), assessing how long the brand has been established (positive factor); (3) market volatility (10%), accounting for the risk of new technology and low entry barriers (negative factor); (4) reach (25%), or the geographic spread of brand sales (positive factor); (5) trend (10%), capturing the upward/downward trajectory of the brand; (6) support (10%), assessing the consistency of marketing support (positive factor); and (7) legal protection (5%), dealing with the firm’s problems in protecting the brand name across markets (negative factor). These factors are combined according to the percentage weights in a brand strength index, which is used to derive a discount factor for projected future earnings. A strong brand with a high index score will yield strong future earnings and thus have a small discount rate (around the 5% typical of a low-risk investment). A weaker brand will have a higher discount rate, reflecting the higher risk associated with its future earnings. The resulting net present value for each brand produces its Interbrand score and associated ranking.


---

15 That said, note that in this research, the data only address one specific aspect of the two measures: how they fared during the 2008 financial market collapse. We are reluctant to make general statements on the overall usefulness of the two metrics (in either absolute or relative terms), as both have a demonstrated value in several other respects.
Appendix B. Equitrend value calculations

The EquiTrend brand equity score is a consumer survey measure that has been collected annually since 1989 for a representative selection of brands in the U.S. market. In 2008, over 20,000 U.S. consumers were surveyed online, and the total number of brands rated was 1170 (each brand received approximately 1000 ratings). In the EquiTrend methodology, the data are weighted to be representative of the entire U.S. population of consumers ages 15 and over on the basis of age by sex, education, race/ethnicity, region, and income. A brand’s equity score is determined by first combining Familiarity, Quality, and Purchase Intent ratings at the individual respondent level. The brand’s total equity score is then aggregated across all respondents with some familiarity with the brand, and the result is indexed on 100. As with Interbrand, the top scores are publicly disseminated, but the brands in the top lists do not overlap consistently. EquiTrend scores for the complete set of brands are available commercially.

Source: www.harrisableinteractive.com

Appendix C

Following the two-step Fama and MacBeth (1973) method, we first estimated the three-factor Fama–French regression model on the four-month period from May 1 to Aug 31, 2008 immediately preceding our analysis period (see Fig. 1). This resulted in 85 daily observations for each brand. The regressions were run separately for each of the 50 brands.

\[
R_{it} - R_{f,t} = \alpha_i + b_1(R_{mt} - R_{f,t}) + s_iSMB_t + h_iHML_t + \epsilon_{it}
\]

where

- \(R_{it}\): stock return for brand \(i\) on day \(t\),
- \(R_{f,t}\): the risk-free rate of return on day \(t\),
- \(R_{mt}\): the market rate of return on day \(t\),
- \(SMB_t\): the difference in returns for small versus big stocks on day \(t\),
- \(HML_t\): the difference in returns for high versus low book-to-market stocks on day \(t\),
- \(i\) = 1, 2, 3, ..., 50 brands,
- \(t\) = 1, 2, 3, ..., 86 trading days.

This error-term has the following properties:

\[E(\epsilon_{it}) = 0; \quad E(\epsilon_{it}^2) = \sigma^2; \quad \text{and } E(\epsilon_{it}\epsilon_{j,t-1}) \neq 0\]

for some finite time period \(t\), indicating possible serial correlation. To correct for serial correlation, we used the Newey–West estimator (Newey & West, 1987) to obtain autocorrelation-consistent estimates. We used the automatic lag selection approach suggested by Newey and West (1994), which led to a lag-length of 3 periods (days) resulting in 83 usable daily observations for the period under consideration.

In the next step, we estimated the following cross-sectional model of 50 brands for Fall 2008 (i.e., Sep. 1–Dec 31, 2008):

\[
DP_i = \beta_0 + \beta_1(b_i) + \beta_2(s_i) + \beta_3(h_i) + \beta_4(t_i) + \eta_i
\]

where

- \(DP_i\): the percentage change in share price for brand during Fall 2008;
- \(b_i, s_i\) and \(h_i\) are the coefficient estimates from Eq. (1);
- \(b_i\): brand value for brand \(i\) (Interbrand or EquiTrend score);
- \(t_i\): error term for brand \(i\).

Finally, we also estimated extended versions of the model in Eq. (2) using additional control variables, including Age, Leverage, Credit rating, and Diversification (as described in the Methodology section).

References


Consumer evaluation of copycat brands: The effect of imitation type

Femke van Horen a,⁎, Rik Pieters b,1

a University of Cologne, Department of Social Psychology, Richard-Strauss-Strasse 2, 50931, Cologne, Germany
b Tilburg Institute of Behavioral and Economics Research (TIBER), Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

A R T I C L E   I N F O
Article history:
First received in 4, November 2010 and was under review for 4½ months
Available online 23 June 2012

Area Editor: John H. Roberts

Keywords:
Copycat brands
Type of imitation
Similarity
Trademark infringement

A B S T R A C T

Copycat brands imitate the trade dress of a leader brand to free ride on the latter’s equity. Copycats can imitate the distinctive perceptual features of the leader brand, such as the lilac color of the Milka chocolate brand, or they can imitate the underlying meaning or theme of the leader brand, such as the “freshness of Alpine milk” theme in Milka. Marketing research and trademark law has focused primarily on the effects of feature imitation. In three studies, the authors demonstrate the success of theme imitation: Consumers consider feature imitation to be unacceptable and unfair, which causes reactance toward the copycat brand. Yet, even though consumers are aware of the use of theme imitation, it is perceived to be more acceptable and less unfair, which helps copycat evaluation.

1. Introduction

Copycat brands imitate the trade-dress of a leading brand, such as its brand name or its package design, to take advantage of the latter’s reputation and marketing efforts. Copycatting is pervasive. For example, Sayman, Hoch, and Raju (2002) observed that blatant package imitation occurred in one-third of the 75 consumer packaged goods categories that they studied. Likewise, in a United States survey, Scott-Morton and Zettelmeyer (2004) found that half of the store brands surveyed were similar to a national brand package in color, size, and shape. Most copycats imitate distinctive perceptual features of the leader brand, such as the color, depicted objects, and/or shape of the package or the letters and sounds of the brand name (Planet Retail, 2007). Thus, copycats imitate the lilac color of the Milka chocolate brand, the bull of the Red Bull energy drink, the spike-shaped bottle of Scope mouthwash,2 the specific letters of the Godiva chocolate brand name, as in “Dogiva”,3 or the Wal-Mart sound, as in Wumart.4

Feature imitation is a strategy that is often used to copy successful leader brands. This type of imitation has received most attention in the marketing and trademark literature (Finch, 1996; Howard, Kerin, & Gengler, 2000; Kapferer, 1995; Loken, Ross, & Hinkle, 1986; Miaoulis & d’Amato, 1978; Zaichkowsky, 2006). Extant research has examined the confusion of copycat brands with leader brands due to various degrees of feature imitation (Howard et al., 2000; Loken et al., 1986) and has investigated the influence of the degree of feature imitation on copycat evaluation (e.g., Van Horen & Pieters, 2012; Warlop & Alba, 2004).

Copycats, however, also use a strategy in which they imitate the underlying meaning or theme of a leader brand, such as the “wildcat” theme of the Puma sports brand, the “freshness of Alpine milk” theme of the Milka brand, or the “traditional, family-produced olive oil” theme of the Bertolli brand. Whereas in feature imitation the focus is on the imitation of one or more of the distinctive perceptual features of the leader brand, in what we term “theme imitation” the focus is on the imitation of the semantic meaning or inferred attribute(s) of the leader brand. To our knowledge, the present research is the first to examine how these different types of imitation influence consumer evaluation of copycat brands.

Theme imitation has received much less attention than feature imitation in the marketing and trademark literature. For example, thirteen of the seventeen cases of copycatting that Zaichkowsky (2006, Chapter 4) documents in her analysis of trademark infringement address feature imitation, while only four cases address theme imitation. One reason for the emphasis on feature imitation might be that feature imitation is easier to detect and prosecute in a court of law; this, of course, does not imply that theme imitation is less effective. The present studies test the hypothesis that imitating the underlying meaning or theme of a

---

leader brand may be a strategy that is more effective than feature imitation. This hypothesis incorporates the idea that when an underlying meaning or theme is imitated, it is likely to be perceived as more acceptable and less unfair than feature imitation because a meaning or theme activates diffuse associations that are not solely linked to the imitated brand. Feature imitations, on the other hand, imitate distinctive perceptual features that belong uniquely to the leader brand and are directly related to the leader brand. Such an imitation strategy is likely to be perceived as unacceptable and unfair and is in turn likely to cause reactance. The three studies described in this paper, which involve different product categories and different verbal (brand names) and pictorial (brand package) imitations provide support for this idea.

1.1. Imitation types

An important precondition for brand imitation strategies to be effective is similarity to the leader brand. To make the leader brand relevant for the evaluation of the copycat, a connection or relation is required. Only then can transfer of knowledge and affect take place (Fazio, 1986). When knowledge of the leader brand is activated and is transferred to the copycat, similarity in the appearance of the brands is generalized to similarity in product quality, thus improving consumers’ evaluation of the copycat (Finch, 1996; Loken et al., 1986).

Copycats most often imitate the distinctive perceptual features of leader brands (visual characteristics, text, sounds), thus showing a type of literal similarity to the leader brand (Gentner, 1983). In simple situations, one might gauge literal similarity between two objects by determining the extent to which they have common and unique features (Tversky, 1977). Thus, the hypothetical brands “Orme” and “Omer” are more similar than the brands “Orme” and “Osve” because the former share all four letters, whereas the latter share only two letters. It is this literal similarity on which most court cases dealing with intellectual property are based (e.g., Adidas Salomon AG vs. Scapa Sports, 2007; Mitchell & Kearney, 2002; Unilever N.V. vs. Albert Heijn B.V., 2005).

However, besides being literally similar through direct imitation of distinctive perceptual features such as letters, colors, shapes, and sounds, two objects can also be semantically similar to each other (Bruce, 1981; Job, Rumiati, & Lotto, 1992). Brands that copy the underlying meaning or theme of other brands aim to make use of the higher-order semantic meanings or inferred attributes of the leader brand. Thus, although the brands “Rome” and “Paris” are semantically similar, they show low literal similarity because they share only one letter, whereas the brand names “Rome” and “Orme” show high literal similarity: they share all four letters but are not semantically similar. In an extreme case, a copycat could even essentially imitate the theme of a leader brand without copying any of the latter’s visual features. Thus, in theme copycatting, the copycat and the leader brand show commonalities with each other not through a display of identical features but instead through the higher-order meaning, theme, or relationship derived from these features.

Of course, themes are displayed through various arrangements of perceptual features. In that sense, theme similarity usually entails at least some level of feature similarity. Therefore, some caution is needed in interpreting the difference between theme copycatting and feature copycatting in an absolute sense. The distinction between similarity in distinctive perceptual features and similarity in higher-order meanings or themes is common in the literature (Gentner, 1983; Gourville & Soman, 2005; Markman & Loewenstein, 2010; Zhang & Markman, 2001).

Feature imitation can occur through imitation of the letters of the leader brand’s name (e.g., by replacing one or more letters of the name or by rearranging them) or through imitation of the distinctive perceptual features of the leader brand’s package design (e.g., the red–white oval logo of Bertolli olive oil or the lilac wrapper of Milka chocolate). Because these distinctive features are exclusively associated with the leader brand, feature imitations are directly linked to the leader brand and will immediately activate a clear representation of the leader brand. Theme imitation can be effected by copying the semantic meaning of the brand name, such as “Spring” (water source) for Sourcy bottled water or by copying the global scene of the package of a leader brand (cows grazing in a meadow in the Alps) for Milka chocolate but presenting it in a visually different way. In contrast to feature imitations, theme imitations are not exclusively associated with the leader brand and will only activate associations that are indirectly linked with the leader brand via a higher-order semantic meaning or an inferred attribute.

1.2. Effectiveness of imitation strategy

Feature copycats are directly linked to the leader brand and almost immediately activate a (positive) image, whereas theme copycats are only indirectly linked to the leader brand. Therefore, one might expect that feature copycats are better able than theme copycats to free-ride effectively on the leader brand’s equity. It is probably this line of reasoning that makes feature imitation a popular copycat strategy. However, based on knowledge accessibility theories (Martin, 1986; Schwarz & Bless, 1992; Wegener & Petty, 1995), we predict differently: we posit that feature imitation will be a less effective imitation strategy than theme imitation.

Research on knowledge accessibility has demonstrated that contextually activated information influences people’s impressions and evaluations of the target (Higgins, 1996; Sherif & Haviland, 1961). The direction of such context effects on assessments of the target can be assimilative or contrastive. Assimilation occurs when evaluation of the target moves toward the contextually activated knowledge, whereas contrast occurs when evaluation moves away from this knowledge. Thus, when compared with luxurious watches like Rolex or Cartier, a moderately luxurious watch may be judged as more or less luxurious.

Various factors determine whether evaluations become more positive or more negative in the vicinity of contextual information (Mussweiler, 2003). One such factor is the perceived appropriateness of the contextually activated information (Martin, 1986; Wegener & Petty, 1995; Wilson & Brekke, 1994). When people are aware that contextual information influences their judgment, they consult their naïve beliefs or theories about the appropriateness of this influence (Petty, Brinol, Tormala, & Wegener, 2007). Such beliefs influence whether people make corrections to their spontaneous judgments (Wegener & Petty, 1995). In the marketplace, consumers are likely to consult their naïve theories of persuasion knowledge when they become aware of the influence of an imitation strategy (Campbell & Kirmani, 2000; Friestad & Wright, 1994). When consumers perceive an imitation strategy to be unacceptable and inappropriate, they tend to correct for the positive feelings induced through similarity.

We predict that consumers will perceive feature copycats as less acceptable and more unfair than theme copycats. Although both types of imitation operate through similarity associations related to the leader brand and make positive knowledge accessible, displays of literal similarity through imitation of the distinctive features of a leader brand are more likely to activate a distinct and clear representation of this brand (“Hey, this looks exactly like X”) because these features are directly linked to the leader brand. Imitation strategies involving literal similarity are therefore likely to be perceived as inappropriate and unacceptable and to cause reactance in consumers, resulting in negative evaluation of the copycat.

Theme imitations, on the other hand, are more implicit and less tangible than feature imitations because the underlying meaning or theme is only indirectly linked to the leader brand. Furthermore, because themes are not only exclusively associated with the imitated brand but also with other objects, brands, or events, such imitation
strategies will be perceived as more acceptable and less unfair than strategies in which distinctive perceptual features are imitated. The evaluative judgment of such copycats should be driven by the affective experiences that are activated by indirect associative links to the leader brand's attributes. These associations are likely to be pleasant and positive because they remind consumers of something they know (i.e., the leader brand), which feels familiar, fluent, and pleasant (Jacoby, Kelley, Brown, & Jasekho, 1989; Moreland & Zajonc, 1982). The positive evaluations associated with the leader brand are likely to become infused into the evaluation of the copycat brand, similar to the way in which affect infusion occurs in other judgments (Forgas, 1995; Schwarz & Clore, 1983). This should cause consumer's evaluation of the copycat brand to move in the direction of the leader brand (assimilation).

1.3. Overview of studies

Three studies were conducted to test the idea that imitation type critically affects copycat evaluation and to probe the psychological process underlying this effect. We predict that theme imitations are evaluated more positively than feature imitations. To establish that theme imitation is indeed a successful imitation strategy, we predict in addition that theme copycats are evaluated more positively than differentiated brands that show no similarity (Hypothesis 1). Furthermore, we predict that theme copycats are evaluated more positively than feature copycats because imitating the underlying meaning or theme is perceived to be more acceptable and less unfair than imitating the distinctive perceptual features (Hypothesis 2).

We tested these predictions using verbal stimuli (brand names, Study 1) and visual stimuli (brand packages, Studies 2 and 3), using a large variety of product categories, and using themes that are either uniquely related to the leader brand (e.g., the wildcat for Puma shoes) or that belong to the whole product category (e.g., the water source for Sourcy bottled water). Studies 1 and 2 examined the basic effect, and Study 3 investigated the underlying mechanism. Empirical support for the hypotheses would imply that copycat evaluation not only depends on how much is imitated (little versus much), about which we know already a great deal, but also on what is imitated (themes or features), about which we currently know very little. Support for the predictions would imply that imitating the underlying theme or meaning communicated by the leader brand is a more successful copycatting strategy than imitating its distinctive perceptual features. This would also have important implications for marketing and trademark law.

2. Study 1: brand names

Study 1 tests the idea that the type of imitation affects copycat evaluation. In this study, brand names were systematically created that either imitated the letters of the leader brand's name (feature imitation), imitated its semantic meaning (theme imitation) or did not show any similarity to the leader brand (no imitation). We used brand names because cases of trademark infringement often address such names. We predict that names that involve theme imitation are evaluated more positively than names that involve feature imitation and names that involve no imitation.

The study controls for two alternative explanations that might account for higher evaluation of theme copycats is that theme brand names have a meaning, whereas feature brand names do not because they simply imitate a letter string. Accordingly, the positive evaluation of theme copycats might be because they adopt a meaningful name rather than a name consisting of a meaningless string. To rule this out, Study 1 includes feature copycat names with meaning. These are brand names that imitate the letters of the leader brand's name and have a meaning (e.g., “Fuma” (smoke) for Puma) but whose meaning does not relate to the semantic meaning (wild cat) of the leader brand (which is imitated by theme copycats). We predict that these copycats are also evaluated more negatively than theme copycats because of their feature similarity with the leader brand.

2.1. Method

2.1.1. Participants and design

Thirty-two paid Dutch undergraduate students (13 females and 19 males, age M = 21.34, SD = 2.04) participated in a 4 (imitation type: no imitation, theme imitation, feature imitation with meaning, feature imitation) × 4 (category: yogurt, bottled water, sport shoes, laundry detergent) within-subject design.

2.1.2. Stimulus development

Four product categories were selected: yogurt, bottled water, sport shoes, and laundry detergent. The leader brands of each of these categories (Almhof, Sourcy, Puma, and Robijn, respectively) are well known in the local market, have strong reputations, and have distinctive brand names. In two of the categories, yogurt and bottled water, the theme communicated by the leader brand was typical for the category as a whole and not unique to the leader brand. The yogurt brand Almhof communicates pastures on farms, and the bottled water brand Sourcy communicates the natural source of the water. In the other two categories, sport shoes and detergent, the theme communicated by the leader brand was unique to the brand. The sport shoe brand Puma communicates the theme of energy and strength symbolized by a leaping wildcat, and the detergent brand Robijn communicates the specialty of a gem (“ruby”).

For each of the four categories, 12 brand names (three names for each imitation type) were developed, resulting in a total of 48 names (see Appendix A). The type of imitation was manipulated in the following way. The brand names in the “no imitation” condition had very low or no similarity to the name of the leader brand (e.g., “Dereon” for Puma, “Imeko” for Almhof). In the “theme imitation” condition, the underlying meaning or theme of the leader brand was imitated; for instance, “Jaguar” for “Puma” (“both wild cats”) and “Weidehove” for Almhof (both meaning “pasture on a farm”). In the “feature imitation” condition, the exact letters of the leader brand’s name were imitated; for instance, “Pumo” for Puma and “Elmhofer” for Almhof. Finally, in the “feature imitation with meaning” condition, the exact letters of the brand name were imitated; in addition, the brand name had a meaning, such as Fuma (smoke) for Puma or “Palmhoff” (courtyard with palms) for Almhof.

A pre-test was conducted to rule out a simple valence explanation for the effects on evaluation in the main study. Participants (N = 15, within-subjects, none participating in the main study) indicated for all 48 brand names whether they considered it a bad brand name (independent of the product it was associated with) on a nine-point scale ranging from 1 (very bad) to 9 (very good). Four brand names (one from each category) showed very low reliabilities and were dropped, leaving 44 brand names for further analyses. An ANOVA with repeated measures revealed no main effect of imitation type, F(3, 42) = 1.38, p = .26, η² = .09, which is desirable. There was a main effect of category, F(3, 42) = 3.76, p = .02, η² = .21, which was qualified by an interaction between category and imitation type. See www.darts-ip.com, last accessed June 2012.
Further analyses of the evaluation of brand names within each of the categories revealed that valence did not differ between imitation types for the categories yogurt, sport shoes, and detergent (all \(p > .05\)). For the bottled water category, the main effect of imitation type was significant, \(F(3, 42) = 5.03, p = .01, \eta_p^2 = .26\). The valence of the feature brand names with meaning was rated higher (\(M = 5.69\)) than the valence of the three other imitation types (\(M_{\text{Diff}} = 4.20, M_{\text{theme}} = 4.73, M_{\text{features}} = 4.90, p < .05\)). We decided not to replace these brand names because there is a limit on brand names that share many letters with the imitated leader brand and they are meaningful. If anything, this result works against our hypothesis because we predict that feature brand names with meaning will be evaluated less positively than theme brand names instead of more positively.

2.1.3. Procedure and measures

Upon arrival, participants were seated in a cubicle in front of a computer. They were told that new products continuously enter the market and that brand names must be developed for them. They were asked to evaluate 44 new brand names within four different categories on a 9-point scale from 1 (negative) to 9 (positive). The product categories and the brand names within each category were randomly presented. Next, as a manipulation check, participants indicated the extent to which they thought that the theme communicated by each of the imitated leader brands was 1 (unique to the brand) to 9 (belonged to the whole category). Finally, as control variables, the participants reported on two items concerning their familiarity with the four leader brands (familiarity with the brand in general and as a product they buy) on a scale ranging from 1 (not familiar at all) to 9 (highly familiar).

2.2. Results and discussion

The manipulation check confirmed that the themes communicated by the imitated leader brands in the product categories sport shoes (wildcat) and detergent (gem) were more unique and did not fit the entire category (\(M_{\text{shoes}} = 4.47, SD = 2.94; M_{\text{detergent}} = 3.56, SD = 2.72\)) than the themes of the imitated brands in the product categories yogurt (pasture on a farm) and water (water source) (\(M_{\text{yogurt}} = 5.75, SD = 2.59; M_{\text{water}} = 6.13, SD = 2.79\)). A planned contrast showed that the difference between the two unique themes combined and the two category-overlapping themes combined was significant, \(F(1, 31) = 14.68, p = .001, \eta_p^2 = .32\).

One differentiated brand name (“Lakai”) in the sport shoes category was dropped because it correlated negatively with the other two brand names in the category, perhaps because of its resemblance to a “servant.” A 4 (imitation type) \(\times\) 4 (product category) ANOVA with repeated measures revealed a main effect of imitation type, \(F(3, 93) = 30.71, p < .001, \eta_p^2 = .50\), of product category, \(F(3, 93) = 13.37, p < .001, \eta_p^2 = .30\), and an interaction between imitation type and category, \(F(9, 279) = 5.82, p < .001, \eta_p^2 = .16\). Table 1 presents the mean evaluation and standard deviations of imitation type within the four product categories.

Table 1

<table>
<thead>
<tr>
<th></th>
<th>No imitation</th>
<th>Theme copycat</th>
<th>Feature copycat with meaning</th>
<th>Feature copycat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yogurt</td>
<td>3.22 (1.47)(^a)</td>
<td>5.53 (1.32)(^b)</td>
<td>3.57 (1.31)(^a)</td>
<td>4.29 (1.51)(^b)</td>
</tr>
<tr>
<td>Bottled water</td>
<td>3.58 (1.17)(^a)</td>
<td>6.33 (1.28)(^b)</td>
<td>4.99 (1.73)(^a)</td>
<td>4.44 (1.78)(^b)</td>
</tr>
<tr>
<td>Sport shoes</td>
<td>4.44 (1.48)(^a)</td>
<td>6.26 (1.22)(^b)</td>
<td>4.64 (1.40)(^a)</td>
<td>4.03 (1.56)(^b)</td>
</tr>
<tr>
<td>Detergent</td>
<td>4.02 (1.48)(^a)</td>
<td>5.22 (1.73)(^b)</td>
<td>3.67 (1.51)(^a)</td>
<td>3.91 (1.79)(^b)</td>
</tr>
</tbody>
</table>

Further analyses of the evaluation of brand names within each of the categories revealed that, as predicted, theme copycats were evaluated more positively than feature copycats (yogurt: \(F(1, 31) = 11.27, p = .002, \eta_p^2 = .27\); bottled water: \(F(1, 31) = 26.31, p < .001, \eta_p^2 = .46\); sport shoes: \(F(1, 31) = 48.00, p < .001, \eta_p^2 = .61\); detergent: \(F(1, 31) = 11.56, p < .002, \eta_p^2 = .27\)). Furthermore, theme copycats were evaluated more positively than feature copycats with meaning (yogurt: \(F(1, 31) = 35.32, p < .001, \eta_p^2 = .53\); bottled water: \(F(1, 31) = 17.89, p < .001, \eta_p^2 = .37\); sport shoes: \(F(1, 31) = 28.08, p < .001, \eta_p^2 = .48\); detergent: \(F(1, 31) = 20.88, p < .001, \eta_p^2 = .40\)) and than differentiated brand names (yogurt: \(F(1, 31) = 73.38, p < .001, \eta_p^2 = .70\); bottled water: \(F(1, 31) = 88.37, p < .001, \eta_p^2 = .74\); sport shoes: \(F(1, 31) = 30.04, p < .001, \eta_p^2 = .49\); detergent: \(F(1, 31) = 14.67, p < .001, \eta_p^2 = .32\)). Follow-up analyses showed that participants were highly familiar with all four leader brands and that familiarity (two measure combined) did not account for the variance in evaluation of the brand names (all Fs < 1, except for the product category yogurt, \(F < 2, p = .18\)).

These results support the hypothesis that theme copycats are evaluated more positively than feature copycats and brand names that do not show any similarity to the leader brand. The findings are consistent across four different product categories, lending importance to their significance. In addition, the demonstration of an effect over four different categories shows that theme imitation is effective even when the theme is unique to the leader brand and not just when the theme is associated with the category as a whole. The results further demonstrate that feature imitations with meaning were evaluated as negatively as feature imitations. This result rules out the possibility that the more positive evaluation of theme copycats as compared to feature copycats is because the former have “semantic meaning” whereas the latter have no meaning.

3. Study 2: package design

Study 1 demonstrates that imitation type affects copycat evaluation. To have high control over the manipulation of different types of imitation, it used brand names. To extend the results from textual stimuli to richer stimuli containing pictorial and textual information, Study 2 used professionally developed images of brand packages. Study 2 establishes the generalizability of the benefits of theme imitation across such richer stimuli and across other product categories. Furthermore, it tests whether the findings on evaluation generalize to consumers’ buying intentions.

3.1. Method

3.1.1. Participants and design

One hundred and thirty-three paid Dutch undergraduate students (80 males and 53 females, age \(M = 22.02, SD = 3.67\)) were randomly assigned to one condition of a three-group design (imitation type: no imitation, theme imitation, feature imitation).

3.1.2. Stimulus development

Images of three packages in the product category “spreadable butter” were designed, with Bertolli as the leader brand (see Fig. 1). Bertolli was the first brand to introduce olive oil in a spreadable form and is the leader in the local market. Bertolli’s trade-dress is unique in the category and consists of both distinctive features (red and white shield, letter type, color) and a clear theme (traditional, family-produced olive oil). To create the feature copycat, the distinctive and unique attributes of the Bertolli trade-dress (the red and white shield, the classical Roman letter type, and the earthly-golden background color) were imitated. As in Study 1, the brand name “Penetoli” shared many of the letters of the leader brand’s name. The theme copycat was created by imitation of the global scene displayed on the Bertolli package (a Tuscan farm on a hill surrounded by pine trees) in a visually different way (see also Miceli & Pieters, 2010). The brand name “Mediterrane” was chosen because this adopts the leader’s main theme. In addition, a no-imitation product.
“Olive Grove”, that did not show any similarity to the leader brand except for the key ingredient, was created.

To test the perceived attractiveness of the packages and to ensure that the packages depicted the intended imitation type, two pre-tests were conducted. The ANOVA of the first pre-test (N = 42, between-subjects, none participating in the main study) revealed that the three packages did not differ from each other in attractiveness (M_Diff = 4.15, SD = 1.14; M_theme = 4.68, SD = 1.28; M_feature = 5.00, SD = 1.15, 7-point scale), F(2, 39) = 1.77, p = .18, r² = .08. Although this test did not reach significance, the means do differ, pointing to a higher attractiveness of the feature copycat, which was not intended. However, a higher attractiveness of the feature copycat would work against our hypothesis (we predict that feature copycats will be evaluated less positively rather than more positively). The second pre-test (N = 15, within-subject, none participating in the main study) ensured that the packages depicted the intended imitation type (Miceli & Pieters, 2010). After the participants had read a description of the two imitation types and the differentiated product, 100% indicated correctly that the package of the feature copycat had read a description of the two imitation types and the differentiated product as a theme copycat). The results of these pre-tests indicate that the design of the copycat brands was intended to be that. In addition, 93% of the participants did this for the differentiated product (7% categorized the differentiated product as a theme copycat). The results of these pre-tests indicate that the design of the copycat brands was successful.

3.1.3. Procedure and measures

The participants were seated in a cubicle in front of a computer and told that the aim of the study was to evaluate a product in the product category “spreadable butter”. Because package designs of supermarket products tend to change regularly, participants were briefly presented with several brands in the product category “spreadable margarine”, including the Bertolli brand, to ensure that the Bertolli package was equally familiar to all participants. Next, one of the three packages (differentiated product, theme copycat, or feature copycat) was displayed on the computer screen for 4 s, and participants were asked to evaluate the product on four semantic differentials with nine-point response alternatives (negative–positive, unattractive–attractive, bad–good, uninteresting–interesting; aggregated evaluation scale, α = .96). Participants then indicated their willingness to buy the product, rating it from 1 (definitely not) to 9 (definitely yes). Finally, participants rated their familiarity with the leader brand on a scale of 1 (not familiar at all) to 9 (highly familiar), their leader brand usage on a scale of 1 (never) to 9 (often), and their evaluation of the leader brand on a scale of 1 (negative) to 9 (positive).

3.2. Results and discussion

3.2.1. Evaluation

ANOVA revealed a significant main effect of imitation type on evaluation, F(2, 130) = 21.78, p < .001, r² = .25. As predicted and consistent with Study 1, planned contrasts showed that the theme copycat was evaluated significantly more positively (M = 6.59, SD = 1.54) than the feature copycat (M = 4.03, SD = 2.21), F(1, 130) = 41.89, p < .001, r² = .24 and the differentiated product (M = 4.88, SD = 1.70), F(1, 130) = 19.55, p < .001, r² = .13. The results further showed that the feature copycat was evaluated more negatively than the differentiated product, F(1, 130) = 4.81, p = .03, r² = .046.

3.2.2. Willingness to buy

A similar pattern of results to that for evaluation was found for willingness to buy. There was a significant main effect of imitation type on willingness to buy, F(2, 130) = 15.05, p < .001, r² = .19. Planned contrasts revealed that participants were more willing to buy the theme copycat (M = 6.63, SD = 1.89) than the feature copycat (M = 4.19, SD = 2.53), F(1, 130) = 26.39, p < .001, r² = .17 and the differentiated product (M = 4.64, SD = 2.15), F(1, 130) = 18.30, p < .001, r² = .12. The difference between the feature copycat and the differentiated product was not significant, F(1, 130) = .95, p = .33, r² = .01.

Follow-up analyses for both evaluation and willingness to buy showed that the control variables did not account for any of the

---

6 We replicated the same pattern of results in the product category milk chocolate (see Study 3 for a precise description of the stimuli). The results (N = 57) demonstrated a significant main effect of imitation type, F(2, 54) = 4.207, p = .02, r² = .14, and planned contrasts showed that, as predicted, the theme copycat was evaluated significantly more positively (M = 5.55, SD = 1.46) than the feature copycat (M = 4.57, SD = 1.24), F(1, 54) = 5.07, p = .03, r² = .09 and the differentiated product (M = 4.40, SD = 1.27), F(1, 54) = 7.34, p = .01, r² = .12.
4.1.1. Participants and design

Participants and design

4.1. Method

when making choices at point-of-purchase. Consumers are often confronted with a number of products showing different attributes, which familiar, insincere, unfair, manipulative (reverse-coded), insincere–sincere and unfair–fair (all nine-point scales) (Brown & Krishna, 2004; Campbell & Kirmani, 2000). The items were aggregated into one acceptability measure (α = .82). The same control variables were included as in Study 2.

4.1.2. Stimulus development

Three packages were created within the product category “milk chocolate”. Milka chocolate was selected as the leader brand because it is well known and its trade-dress is unique both in its features and in its theme. To create the theme copycat (“Montana”), the global scene displayed on the package of the Milka brand (cows grazing in the fresh, nutritious fields of the Alps) was imitated in a visually different way (see Fig. 2). To create the feature copycat (“Lecha”), the distinctive features of the Milka brand (the lilac color, the Milka cow, and the creamy font type) were copied. In addition, a differentiated product (“Davinia”) was created.

A first pre-test (N = 45, between-subjects) demonstrated that the three packages did not differ in attractiveness (Mdiff = 4.13, SD = 1.41; Mtheme = 4.27, SD = 1.53; Mfeature = 4.53, SD = 1.30, F(2, 42) = .31, p = .74, η2p = .015), and a second pretest (N = 15) established that the manipulation of imitation type was successful (100% correct categorization for feature and differentiated and 87% correct categorization for theme (13% categorized it as a feature copycat)).

4.1.3. Procedure and measures

The general set-up was similar to that used in Study 2, the only difference being that, instead of just one product, all three products were presented on the screen for several seconds. The participants were then asked to evaluate the products (α > .96), to indicate their willingness to buy, and to choose one of the products. To test the extent to which familiar, fluent, and pleasant affective experiences were activated by each of the imitation types, participants were asked to indicate to what extent the product ‘feels good’ and ‘feels familiar’ ranging from 1 (not at all) to 9 (very much). Furthermore, to assess the acceptability of the imitation strategy, participants were asked to indicate the degree to which they thought that the similarity of the product with the Milka brand was unacceptable–acceptable, not manipulative–manipulative (reverse-coded), insincere–sincere and unfair–fair (all nine-point scales) (Brown & Krishna, 2004; Campbell & Kirmani, 2000). The items were aggregated into one acceptability measure (α = .82). The same control variables were included as in Study 2.

4. Study 3: acceptability

Studies 1 and 2 established the basic effect (described in Hypothesis 1) using brand names and brand packages. Study 3 probes the underlying mechanism and addresses Hypothesis 2. Because theme imitations are not exclusively associated with the leader brand and are only indirectly linked to the leader brand via higher-order semantic meanings or inferred attributes, we predict that consumers will perceive this type of imitation as more acceptable and less unfair than when the distinctive perceptual features of the leader brand are directly imitated. Study 3 tests the extent to which positive evaluation and higher buying intention of theme copycats compared to feature copycats is mediated by the acceptability of the imitation type. In addition, to assess whether positive evaluation of the theme copycat is indeed the result of the use of an imitation strategy, we test whether familiarity-induced positive effects account for the more positive evaluation of the theme copycat compared to the differentiated product. Study 3 also tests whether effects on evaluation transfer to choice. To test the effects on choice, a within-subject instead of a between-subjects design was used. The use of a within-subject design is also closer to reality because consumers are often confronted with a number of products showing different imitation types that are simultaneously displayed on the shelves when making choices at point-of-purchase.

4.1 Method

4.1.1. Participants and design

One-hundred and six paid Dutch undergraduate students (55 males and 51 females, age M = 23.08, SD = 3.97) were randomly allocated to a condition of a 3 (imitation type: no imitation, theme imitation, feature imitation) × 3 (presentation order) mixed design with imitation type as the within-subject factor and presentation order as the between-subjects factor. The presentation order was varied across participants to control for order effects.

![Fig. 2. Stimuli used in Study 3.](image-url)
4.2. Results and discussion

There was no effect of presentation order on any of the dependent measures; therefore, this factor was excluded from further analyses. The evaluation and willingness to buy measures were highly correlated (rs > .84) and collapsed into a single measure.

4.2.1. Evaluation

Repeated measures ANOVA revealed a main effect of imitation type, F(2, 210) = 12.63, p < .001, ηp² = .11. Consistent with the results of Studies 1 and 2, planned contrasts showed that the theme copycat (M = 5.93, SD = 1.73) was evaluated more positively than the feature copycat (M = 4.62, SD = 1.92), F(1, 105) = 41.80, p < .001, ηp² = .29 and the differentiated product (M = 5.01, SD = 2.05), F(1, 105) = 10.32, p = .002, ηp² = .09. The evaluation of the feature copycat did not differ from the evaluation of the differentiated product F(1, 105) = 1.70, p = .20, ηp² = .02. Follow-up analyses showed that the control variables did not account for any of the variance (all Fs < 1.5).

4.2.2. Choice

Frequency analysis demonstrated that the theme copycat was more frequently chosen (49%) than the feature copycat (18%) or the differentiated product (33%). Additional conditional log regression analysis showed that the differences in probability of choice were significant (theme versus feature: β = −1.01, z(318) = −3.76, p < .001, theme versus differentiated: β = −.40, z(318) = −1.81, p = .07, and differentiated versus feature: β = −.61, z(318) = −2.14, p = .03).

4.2.3. Mediation analysis

The first mediation analysis tested whether positive evaluation of the theme copycat compared to the feature copycat is mediated by acceptability of the imitation strategy. A regression analysis was conducted with two orthogonally contrast-coded variables (i) differentiated versus theme/feature and (ii) theme versus feature) and acceptability as predictors (Preacher & Hayes, 2008). The results demonstrated that the contrast variable feature (−.5) versus theme imitation (.5) predicted the dependent variable evaluation and the mediating variable acceptability significantly (β = 1.31, t = 4.99, p < .001, β = 2.74, t = 12.13, p < .001, respectively). When the mediating variable was included in the full model, the results revealed a significant effect of the mediating variable on evaluation (β = .20, t = 3.04, p = .003) whereas the significance of the contrast variable dropped (β = .77, t = 2.46, p = .015). Bootstrapping analyses revealed that acceptability indeed mediated the positive evaluation of the theme copycat compared to the feature copycat, yielding a point estimate for the indirect effect of 1.08 and a 95% confidence interval of .69 to 1.49.

4.2.4. Follow-up study

The results indicate that theme imitations are evaluated more positively than feature imitations because consumers perceive theme copycats as more acceptable and less unfair than feature copycats. The question remains, however, whether consumers perceive theme imitations as acceptable because they are unaware of the imitation strategy or whether they are aware of the used imitation strategy but nevertheless perceive such imitations as more fair and acceptable.

A follow-up study (N = 55, between-subjects) provided more insight into the exact process. The same milk chocolate packages were used as in the main study. Participants were asked to indicate on a scale from 1 (definitely not) to 9 (totally) whether it was clear to them that the manufacturer of this product made use of an imitation strategy. Participants were then asked to indicate on four semantic differentials (all nine-point scales) to what extent they thought the specific imitation strategy used by the manufacturer was acceptable, fair, and allowed (aggregated acceptability measure, α = .96). Finally, they were asked to indicate how much the particular package design made them think of the Milka chocolate package using a scale ranging from 1 (not at all) to 9 (very much).

The results demonstrated that people were more aware of the imitation tactics of the theme copycat (M = 7.78, SD = .94) than of the differentiated product (M = 4.74, SD = 2.38), F(1, 52) = 36.20, p < .001, ηp² = .41 and that they were marginally less aware of the strategy of the theme copycat than that of the feature copycat (M = 8.67, SD = .59), F(1, 52) = 3.01, p = .09, ηp² = .06. Despite this awareness, the imitation strategy used by the theme copycat was perceived as more acceptable and fair (M = 5.00, SD = 2.31) than the strategy used by the feature copycat (M = 2.61, SD = 1.26), F(1, 52) = 16.01, p < .001, ηp² = .24, whereas the difference in acceptability between the theme copycat and the differentiated product (M = 5.89, SD = 1.65) was not statistically significant F(1, 52) = 2.31, p = .14, ηp² = .04. In addition, the package of the theme copycat made people think of the Milka chocolate package (M = 8.39, SD = .78) as much as the package of the feature copycat did (M = 8.39, SD = 1.46), F(1, 52) = .00, p = 1.00, ηp² = .00. This indicates that the theme copycat was indeed perceived as a copycat.

The results of Study 3 replicate the findings of Studies 1 and 2. In addition, they demonstrate that, besides evaluation and willingness to buy, imitation type also affects copycat choice: consumers chose theme copycats more often than feature copycats and differentiated products. Furthermore, they reveal that theme imitation is evaluated more positively than feature imitation because theme imitation is perceived to be more acceptable and less unfair, despite consumers’ awareness of the imitation strategy used.

5. General discussion

In the marketplace, consumers can be confronted with two types of imitation: feature imitation and theme imitation. Building on knowledge accessibility theories, we propose that these two imitation strategies affect copycat success differently. The three studies described in this article demonstrate that imitation of the underlying meaning or theme of a design is a more effective copycatting strategy than imitation of its distinctive perceptual features. This is important because feature imitation has received the most emphasis in the marketing and trademark literature. In Studies 1 and 2, we establish the basic effect and show that theme

7 In the current design, awareness of imitation tactics was prompted. This may sometimes occur in practice, e.g., when consumers look for similarities, but it is not very typical. Despite this explicit awareness, the results showed that the imitation tactics employed are still judged to be acceptable and fair. It is unlikely that these results will be different when people are aware without being prompted.
imitation is more effective both when the theme is unique to a leader brand and when it is associated with the category as a whole. We show that this effect generalizes across product categories and transfers from evaluation to buying intentions and choice. Study 3 demonstrates that acceptability of the imitation tactic is an important psychological process that underlies the basic effect.

As in other research on copycatting, the specific manipulations of feature copycats used in the present studies depend on the features used by the existing leader brands in the market; therefore, the appearance of the copycat relies strongly on features that are distinctive for the imitated leader brand (e.g., a specific type font or color). Because the type font was particularly distinctive in the package designs of the Milka and Bertolli brands, one may argue that it is the use of a specific font that accounts for the current results rather than feature imitation in general. However, our consistent findings using brand names in Study 1, where features and themes were manipulated merely by changing or rearranging letters, demonstrate the generality of the effects.

These findings have implications for marketing theory and practice. First, to our knowledge, the present studies are the first to demonstrate that imitation type influences consumer evaluation of copycat brands. Because the majority of copycats are feature-based, the marketing literature has focused primarily on consumer responses to this type of imitation (Howard et al., 2000; Kapferer, 1995; Loken et al., 1986) and on the effect of the degree of imitation on evaluation (Miaoulis & d’Amato, 1978; Van Horen & Pieters, 2012; Warlop & Alba, 2004). The current research shows that, in addition to how much is imitated, what is imitated (theme versus features) is of crucial importance as well.

The second feature of the current studies is that they build on and extend previous research on knowledge accessibility effects in essential ways. In our studies, the standard (leader brand) against which the target (copycat brand) was evaluated was activated by the target itself rather than being induced by contextual information as in most previous research (e.g., Mussweiler, 2003). Thus, our studies provide a strong test of the effects of imitation type on consumer evaluation. In addition, instead of manipulating the standard (e.g., an extreme versus a moderate standard; Smeesters, Mussweiler, & Mandel, 2010), our studies kept the standard (the leader brand) constant.

A third significant implication of our results is that they challenge the prevailing idea in trademark legislation that copycats that imitate the distinctive perceptual features of the leader brand are the most perilous and require the most attention from the imitated leader brands. This idea is reflected in a recent decision of the European Court that only highly distinctive, but not less distinctive, features of brands can receive protection (Digipos Store Solutions vs. Digi International, 2008). To explore this issue further, we surveyed a sample of 44 lawyers specializing in intellectual property at a trademark law conference in The Netherlands. The lawyers were asked to indicate the extent to which they believed a consumer would positively evaluate a product that imitated the highly distinctive perceptual features of a leader brand and a product that imitated several less distinctive features. (Both products were evaluated on a nine-point scale ranging from 1, definitely not, to 9, definitely yes). Indeed, lawyers believed that consumers would evaluate copycats that imitated the highly distinctive features more positively ($M = 7.02, SD = 1.85$) than copycats that imitated several less distinctive features ($M = 5.50, SD = 1.55$), $F(1, 43) = 15.84, p < .001$, $\eta^2 = .27$. The current studies, however, demonstrate that the converse is true. This result suggests that (at least Dutch) trademark lawyers might be well advised to focus less on highly distinctive feature copycats because consumers evaluate them positively.

For manufacturers of copycats, it may be worthwhile to consider investing in brand names or package designs that imitate the underlying meanings or themes of leader brands rather than their distinctive features. For manufacturers of leader brands, on the other hand, it is advisable to invest principally in the distinctive features of the package. Investing in visually unique package designs is not only important in enabling consumers to distinguish leader brands from other brands in a cluttered environment (Van der Lans, Pieters, & Wedel, 2008) and to facilitate brand recognition and recall, but, as the current research suggests, is also a powerful tool in warding off imitation attempts by other brands. For manufacturers of leader brands that are being “theme copied,” the present findings provide suggestions for tests and/or metrics to establish that theme imitation is occurring and that their brand equity is being unfairly hijacked by copycat brands.

A question for future research concerns the exact roles that type and degree of similarity play in copycat evaluation. That is, do consumers prefer theme imitations simply because they are less similar than feature imitations and dislike feature imitations because they are too close to the leader brand (Carpenter & Nakamoto, 1989)? Do feature imitations tend to be evaluated more positively when they imitate the leader brand to a lesser extent? Due to the abstract, implicit character of theme imitations, it might indeed be the case that theme imitations are perceived as less similar than feature imitations. On the other hand, copycats that imitate a theme that is highly unique to a specific leader brand, might be perceived as highly similar as well. When all detergents but one depict soft objects on their packaging (e.g., a bear, duck, or baby) and one detergent displays a ruby, displaying a piece of jade is likely to be perceived as highly similar despite the fact that all the distinctive features of the packaging are different. However, we conjecture that even highly similar theme copycats would still be evaluated more positively than highly similar feature copycats. This is because the inferred attribute or semantic meaning imitated by the theme copycat is, in addition to being associated with the leader brand, also associated with other objects, brands, and events; thus, imitation will be perceived as less unfair than when the exact features of a brand are imitated.

In a first attempt to address the issue of the influence of the type and degree of similarity, we conducted follow-up mediation analyses on the data of Study 3. In these analyses, similarity and similarity-squared were added to each of the two models described in Study 3 predicting brand evaluation. The results demonstrated that acceptability still mediated the positive evaluation of theme versus feature ($\beta = .27, p < .001$) and that familiarity-induced affect mediated the positive difference between theme and differentiated product ($\beta = .78, p < .001$), whereas similarity and similarity squared did not mediate the effects in either model ($|\beta_{Sim} = .23, p = .24; \beta_{Sim^2} = -.01, p = .70$ and $|\beta_{Sim} = .18, p = .15; \beta_{Sim^2} = -.02, p = .11$, respectively). This suggests that independent of the degree of similarity, the type of similarity – feature or theme – plays an important role in copycat evaluation.

Another avenue for future research would be investigation of the role of time pressure or lack of processing resources during evaluation and choice. All our studies were conducted under fairly high levels of involvement, and participants had sufficient time to evaluate and choose the products. Perhaps feature imitations compete more effectively under low levels of involvement because correcting initially positive evaluations requires available resources. Future studies could also probe the effect of leader brand presence on the evaluation of theme copycats. The positive evaluation of theme copycats may dwindle when the situation prompts direct comparisons between copycat and leader brands because the advertising strategy used then becomes more apparent. On the other hand, evaluation could be unaffected because the imitated theme is not exclusively associated with the leader brand, and imitation of such themes may still be considered acceptable even when direct comparisons are made.

Recently, the British Brands Group described copycats as products that “hijack distinctive features of a brand’s packaging to trick shoppers into buying something they believe to be the brand” (Shelf Life, 2008).
The present research reveals, however, that consumers may not be easily tricked by explicit hijacking of distinctive perceptual features. Instead, consumers appear to more easily fall prey to the persuasive appeal of copycats when these piggyback on the success of leader brands in a more implicit and less distinctive manner by imitating their themes.

Acknowledgments

The authors wish to thank the Editor, Associate Editor, and two anonymous reviewers for their helpful comments on an earlier version of this paper, and Simon Kee (www.simonkee.nl) for his help in developing the stimuli used in Studies 2 and 3.

Appendix A

Brand names used in Study 1.

<table>
<thead>
<tr>
<th>Product category</th>
<th>Leader brand</th>
<th>Type of imitation</th>
<th>Brand names</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yogurt</td>
<td>Almhof</td>
<td>No imitation</td>
<td>IMEKO, GLANBIA, LANDLIEBE, PALMHOF, ALMHOF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Theme imitation</td>
<td>MEADOW FARM, WEIDHOEVE, LAMHOF, ALMHOF</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feature imitation with meaning</td>
<td>SOURTY, SOURCE</td>
</tr>
<tr>
<td>Bottled water</td>
<td>Source</td>
<td>No imitation</td>
<td>PENTA, UZARSA, SPRING, SOURCE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Theme imitation</td>
<td>DASANI, FOUNT WEL, SOURLY</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feature imitation with meaning</td>
<td>SOURTY, SIERCY, SOURLY</td>
</tr>
<tr>
<td>Sport shoes</td>
<td>Puma</td>
<td>No imitation</td>
<td>AKOO, DEREON, LEOPARD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Theme imitation</td>
<td>LAKAI, PANTHER, TIGER</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feature imitation with meaning</td>
<td>PUMAS, PUMA, SUMMA</td>
</tr>
<tr>
<td>Laundry detergent</td>
<td>Robijn</td>
<td>No imitation</td>
<td>BURI, DROMAL, PUMO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Theme imitation</td>
<td>NIBMA, SMARAGD, SAFIER</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Feature imitation with meaning</td>
<td>(ZUMA), KORABIN</td>
</tr>
</tbody>
</table>

Note: Brand names in brackets were dropped after the pretest. The themes communicated by the leader brands of the product categories sport shoes (Puma: energy symbolized by a leaping wildcat) and laundry detergent (Robijn: specialty of a gem) were unique to the leader brands. The themes communicated by the leader brands in the product categories yogurt (Almhof: pasture on a farm) and bottled water (Source: natural source of the water) were typical for the category as a whole and were not unique to the leader brand.

References

Adidas Salomon AG vs. Scapa Sports (2007), IJN: B87456, Gerechtshof Amsterdam, 1278/06 KG.
Carbonell vs La Espanola-Acetes del Sur vs Kooph (2009), European Court of Justice, IPPT20090903, C-498/07P.
DigiPos Store Solutions vs. Digi International (2008), High Court of Justice (Chancery Division) R.P.C., 125: 591–625.
Planet Retail (July).


Seeing the forest despite the trees: Brand effects on choice uncertainty

Christine Eckert a,⁎, Jordan J. Louviere a, Towhidul Islam b

a Marketing Discipline Group, Centre for the Study of Choice (CenSoC), UTS Business School, University of Technology, Sydney, Australia
b Department of Marketing and Consumer Studies, University of Guelph, Canada

A R T I C L E   I N F O

Article history:
First received in 1, February 2010 and was under review for 7 months
Available online 28 March 2012

Area Editor: Bart J.J.A.M. Bronnenberg

Keywords:
Brand equity
Scale heterogeneity

A B S T R A C T

Prior research on brand equity suggests that consumers use brands as signals to reduce uncertainty and perceived risk. Erdem and Swait (1998) developed a conceptual framework based on information economics and signaling theory to explain how equity is created, maintained and transferred over time that involves seven theoretical constructs. This paper reviews the impact of brand-equity-associated brand utility on the scale of the indirect utility function (i.e., the inverse of the error variance); we argue that higher brand-equity-associated brand utility reduces the need for consumers to review previously formed preferences. We combine a brand utility experiment with a brand feature experiment to estimate the effects of brand-equity-associated brand utility scores on choice. We find that higher brand-equity-associated brand utility leads to higher choice consistency, which can drive increases in market share.

Crown Copyright © 2012 Published by Elsevier B.V. All rights reserved.

1. Introduction

The concept of a brand (defined as “a name, term, sign, symbol or design, or a combination of them which is intended to identify the goods and services of one seller or a group of sellers and to differentiate them from those of competitors”; Kotler, 1997, p. 443) is widely regarded as a key marketing principle. Different research streams focus on different roles that brands play in consumer choices. One stream focuses on the impact of brands on consumer utility in random utility choice models (e.g., Kamakura & Russell, 1993; Louviere & Johnson, 1988; Park & Srinivasan, 1994). A second stream focuses on price premiums that consumers are willing to pay (e.g., Kamakura & Russell, 1993; Park & Srinivasan, 1994) or differences in price sensitivity for strong brands (e.g., Keller, 1993; Sivakumar & Raj, 1997). Other streams focus on ways to measure brand health and brand satisfaction, among other issues (e.g., Ailawadi, Lehmann, & Neslin, 2003; Bloemer & Kasper, 1995).

In this paper, we adopt the perspective of Erdem and Swait (1998), who view brands in an information economics framework. They propose that markets are characterized by imperfect and asymmetric information and that “consumer uncertainty about product attributes may exist even after active information gathering (for experience attributes) or after consumption (for long-term exposure or credence attributes)” (Erdem & Swait, 1998, p. 138). Companies can reduce this uncertainty by sending signals about their product quality, for example, through advertising (Milgrom & Roberts, 1986) or manufacturers’ warranties (Lutz, 1989). In the information economics perspective on brand equity, a brand name itself serves as a signal, and thus, “consumer-based brand equity is defined as the value of a brand signal to consumers” (Erdem & Swait, 1998, p. 133). For example, umbrella branding (in which firms use the brand name of established products for new products) is one way to send a quality signal about a new product to consumers. This perspective on brand equity ascribes costs to information acquisition processes that consumers use to resolve uncertainty. Higher brand equity, which is defined as a strong brand signal, reduces these information costs, in turn leading to higher brand utility (Erdem & Swait, 1998).

Over the past several decades, product markets have become more variable, with a proliferation of SKUs in many product categories. Additionally, and more recently, major supermarket chains across the world have begun to move into private-label products, further increasing the number of choices and diversity available to customers. Not only are there more signals and more diverse signals being sent by firms about an increasing number of brand offers, but the proliferation of these signals now also occurs across many more communication channels, such as Facebook, Twitter and other new media. Moreover, the availability of consumer product ratings via consumer-review websites and e-commerce websites, such as http://www.yelp.com/, http://www.walmart.com/ and http://www.amazon.com/, increases the complexity of brand signals.

A potentially important issue that has been largely ignored in the brand equity literature is the impact of a brand name on choice consistency. Choice consistency is one of several possible unobserved utility components in random utility theory-based choice models.2

⁎ Corresponding author. Tel.: +61 2 95143538.
E-mail addresses: Christine.Eckert@uts.edu.au (C. Eckert), Jordan.Louviere@uts.edu.au (J.J. Louviere), islam@uoguelph.ca (T. Islam).

1 We recognize that there is a trend, particularly in the USA, for some retailers to restrict assortments of products.
2 We are only aware of a study by Swait and Erdem (2007), which showed that one basic component of brand utility, namely, brand credibility, has a positive impact on decision makers’ ability to discriminate between product utilities in choice situations.
That is, the stochastic component of utility (the so-called “error component”) can be decomposed into several possible subcomponents, such as variability in choices due to mistakes, inattention, differences in familiarity with choice options and model misspecification. We propose that brands that provide strong signals to consumers are also likely to exhibit more consistent choices in scanner panel and choice experiment data sources. Strong brands help a “decision maker more strongly discriminate between that brand and others, because the evaluation of the former product may be less subject to idiosyncratic uncertainties (e.g., different levels of knowledge about attributes) compared with that of other brands” (Swait & Erdem, 2002, p. 307). This discrimination further suggests that stronger brands can simplify consumer decision processes and reduce the need to reevaluate products, making purchase decisions easier for consumers.

One measure of choice consistency is the scale of the indirect utility function, which is inversely proportional to the standard deviation of the error component. This scale determines how “in different contexts and for different decision makers, the same systematic utility difference can result in more-extreme choice probabilities” (Swait & Erdem, 2007, p. 682). As Swait and Louviere (1993), Louviere, Hensher, and Swait (2001, Chapter 8), Salisbury and Feinberg (2010) and Fiebig, Keane, Louviere, and Wasi (2010) indicate, if individuals differ in their scales, there can be significant implications for marketing policies that rely on choice modeling results. For example, suppose there are two segments of consumers, one with considerable experience in a category that makes very consistent product choices, and a second that is new to the category and makes much less consistent choices. Even if the consumers in these two segments use identical decision rules (i.e., choice models and associated indirect utility specifications) to make choices, their observed (and predicted) choice probabilities (i.e., proportions) will differ. That is, the segment that makes choices more consistently should exhibit a wider range of choice probabilities in any particular context (e.g., choice set, choice occasion, etc.) than the less consistent segment. Indeed, at the extreme of almost perfect choice consistency, observed choice probabilities will be close to zero and one whereas when choice consistency decreases, observed choice probabilities will be close to $1/J$, where $J$ is the number of choice options offered.

Thus, managers need to better understand that the range of predicted share is likely to be smaller/larger for segments (individuals) with lower/higher error variances or choice consistency, and this should be taken into account when making marketing policy decisions. In turn, this implies that choice consistency or the scale of the utility estimates is likely to be useful for positioning and targeting. Prior literature identified several factors that are associated with differences in the scale of the indirect utility function, including information frames (e.g., Swait & Adamowicz, 2001), labeled alternatives in choice sets (e.g., brand names) and individual differences, such as education/literacy, age and involvement (see, e.g., Louviere et al., 2001, Chapters 8 and 13).

Consequently, the purpose of this paper is to test whether brand names affect consumer choice consistency, which in turn affects consumer choices. We show that brands systematically affect consumer choice consistency in a series of discrete choice experiments, and our results suggest that if managers understand and can predict differences in effects on choices due to “pure preferences” and choice consistency, it should be possible to make more effective use of a brand’s marketing mix so as to affect both preferences and choice consistency. For example, managers can manipulate the strength of a brand signal by carefully choosing suitable advertising and product line extension branding strategies, which send out consistent quality signals to the consumer and thus strengthen a brand’s credibility. Similarly, different advertising and communications media also likely vary in their credibility, and thus, managers can likely improve brand credibility and other equity signals by appropriate media choices, a topic that we leave for future research.

In a practical sense, differences in choice consistency matter in marketing research because, as noted by Fiebig et al. (2010) and Salisbury and Feinberg (2010), if scales differ across contexts, segments and individuals (among other things), these scale differences must be accounted for in traditional choice models (including models that attempt to capture preference heterogeneity) to avoid confounding partworth estimates with these differences. Scales are likely to differ by brand and attribute as well as by the individual, as shown by Louviere and Eagle (2006) and Louviere and Meyer (2007). If the estimated utility of a particular brand is due in part to pure preference for the brand and in part to the choice consistency associated with that brand, existing choice models such as conditional logit or its extensions (i.e., random effects or latent class models) can potentially yield incorrect and misleading results. That is, ignoring error variability differences in choice models may result in biased parameter estimates rather than a simple loss of efficiency (Swait & Erdem, 2002).

This paper tests the hypothesis that brand-utility- or equity-related brand utilities affect unobserved choice consistency. To empirically examine this proposition, we rely on the Erdem and Swait (1998) theory of brand equity in a discrete choice setting using respondents’ choices from two related discrete choice experiments (DCEs) (Louviere & Woodworth, 1983; Street, Burgess, & Louviere, 2005). One DCE asks participants to choose among alternatives described by the Erdem and Swait theory constructs (a brand equity DCE), and a second DCE asks them to choose among choice options described by attributes/features (a traditional DCE). Thus, the brand equity DCE is used to measure brand equity as defined by signaling theory, and we test the effects of the brand equity measures on choice consistency in the second DCE. To determine the magnitude of the choice consistency effect, we use recent advances in random utility models to estimate a more flexible and general model that simultaneously allows for a distribution of scale and a distribution of preferences. In particular, we develop and apply a model that combines the Generalized Multinomial Logit, or G-MNL model (Fiebig et al., 2010), with a variant of the Random Coefficients Generalized Scale Multinomial Logit (RGCSMNL) model by Swait and Erdem (2002).

The remainder of the paper is organized as follows. We first describe the two discrete choice experiments and then discuss conceptual and statistical model considerations, followed by the presentation of results from six empirical datasets. We conclude by discussing the managerial implications of this study and directions for future research.

2. Research approach

2.1. Description of the discrete choice experiments

We designed and implemented two DCEs for each of six different product categories. The categories are cross-country airline flights, car insurance and four electronic products categories that we categorize as follows: a) listening devices, b) visual entertainment devices, c) auditory devices and d) after-market add-on devices to enhance communication. We disguise the last four product categories and their associated features/levels to ensure confidentiality but note that product features are not germane to the tests that we conduct. The airline flights and car insurance data sets were collected with the sole purpose of testing the hypotheses in this paper, but the four other DCEs were designed and implemented as part of a larger project conducted in collaboration with a major electronics producer. Thus, the data from the last four categories lack several features that we outline below.

The two DCEs in each category were as follows: 1) a brand-anchored design (Louviere & Johnson, 1990) in which choice alternatives (hereafter, choice “options”) are described as ’like brand A’ (where “A” is a real brand name) on each of the Erdem and Swait brand equity constructs (see Appendix A.1 for a DCE screenshot of
the airline category) and 2) a product feature (or “attribute”) design where key attributes of real brands were varied (see Appendix A.2 for a DCE screenshot of the airline category).

The purpose of the first DCE was to estimate the brand-equity-associated utilities for each brand for each individual. We then used this measure of brand utility in the second DCE to measure the impact of brand-equity-associated brand utility on choice consistency when respondents chose among complete product offerings.

We randomly assigned 400 respondents to each of the airline and the car insurance categories and collected data from an online panel in April 2010. For the four electronics categories, we randomly assigned 512 respondents to each of the four categories and collected data from an online panel in September 2008. In each category, every respondent completed both DCEs, allowing us to calculate brand-equity-associated brand utility scores for each person from DCE 1 and match these scores with brands in DCE 2. In the airline and car insurance categories, respondents were further randomly assigned to the sequences as follows: a) the brand-anchored DCE first and the product-feature DCE second or b) vice versa. The DCEs may artificially prime brand associations or decompositions of brands with respect to the underlying brand equity constructs, so we used the two orders to test for priming effects in both categories. There were no significant differences (p-value of χ²-test = 0.256) in the question order for the airline choices; however, we found significant differences (p-value of χ²-test = 0.000) in the car insurance category. Because the data sets in the electronics categories were initially not collected for the purpose of this paper, we cannot provide any priming tests for them. We discuss possible implications of this priming at a later stage in this paper.

There are four target brands of interest in both airline and car insurance categories, and eight brands in each of the electronics categories. As noted earlier, DCE 1 is a brand-anchored conjoint experiment with four choice alternatives per choice set. In the electronic categories (for which we collected the data first), each alternative was described by the following seven constructs in the Erdem and Swait (1998) theory of brand equity: consistency, credibility, clarity, investment, risk, quality and search/time cost. The airline and car insurance DCE surveys followed a later study by Erdem and Swait (2004) in using only six of the seven original constructs proposed by Erdem and Swait (1998); that is, the “clarity” construct is excluded. The DCE 1 design used an orthogonal main effects plan for its initial design and then proceeded as in Street et al. (2005) to develop a design in 8 choice sets for the electronics categories. We used the same design approach to generate designs for airlines and car insurance, adding four more choice sets for hold-out validation. Each choice set consisted of four alternatives, and each design was D-optimal under the Street and Burgess (2007) criterion of a true null hypothesis for the model estimates.

We used three statements to describe each construct in DCE 1. We modified previous statements used to test the Erdem and Swait theory in a structural equation framework (e.g., Erdem & Swait, 1998), as shown in Appendix B. The modified statements were tested in several rounds of pilot tests to ensure a large majority of respondents agreed that the statements represented each of the constructs. It is important to note that we did not ask respondents to evaluate each brand on each statement; instead, we described a particular choice option using phrasing such as “like Virgin Blue in consistency”, “like Tiger Airways in quality” and “like Qantas in credibility”. The respondents’ choices jointly reveal the effect of each construct and the position of each brand on each construct, allowing us to estimate the brand-equity-associated utility for each brand. In each choice set, respondents reported their most and least preferred options of the four given.

In the product feature experiment (DCE 2), respondents were again offered four choice options per choice set, described by the same brand names as in DCE 1 with additional product attributes (e.g., flight duration or fare in the airline category). As with the brand-anchored DCE design, we used an orthogonal main effects plan as a starting design to implement the Street et al. (2005) approach to construct a DCE with (again) 8 and 8 + 4 choice sets for the electronics categories and the airline and car insurance categories, respectively. As in DCE 1, respondents indicated their most and least preferred options in each choice set. Unlike DCE 1, we also asked a third question in each choice set, namely, whether a respondent would purchase the option that they indicated as their most preferred choice if it was available at the time of their next purchase. Answers to the third question were not used to conduct the tests of interest in this paper, so we do not consider them further.

2.2. Conceptual framework

The test of interest in this paper is whether brand-equity-associated brand utility scores from DCE 1 affect the scale (i.e., choice consistency) in DCE 2. Conceptually, we expect that the features in DCE 2 will not capture all of the utility associated with the brands offered in DCE 2 due to unobserved attributes and/or omitted variables, model misspecification and so forth. Indeed, to the extent that brands act as signals as suggested by Erdem and Swait (1998), the brand signal information measured in DCE 1 should be able to explain the variability in choices in DCE 2, in which respondents have to retrieve brand information from memory while trading-off brand names with other product attributes. Thus, to the extent that the brand names associated with the constructs or some subset of them systematically explain choices of alternatives in DCE 1, there should be additional information about brand-equity-associated brand utility for each individual. The null hypothesis is that the brand-equity-associated utility measures from DCE 1 are irrelevant and provide no additional systematic information about the choices in DCE 2.

That is, under the null, all brand-relevant information is captured by the brand utilities in the indirect utility functions estimated from DCE 2. Thus, there should be no systematic effects of brand-equity-associated brand utility measures derived for each respondent in DCE 1 on their choices in DCE 2. However, we expect that respondents in DCE 1 who exhibit higher/lower brand-equity-associated utilities for particular brands will also exhibit more/less consistent choices in DCE 2. That is, respondents who consistently choose options in DCE 1 with particular brand-construct levels should have higher scales (i.e., lower error variability = more consistent choices) in DCE 2. In turn, this suggests that we should find a positive interaction between brand-equity-associated brand utility scores from DCE 1 and the summary measures of the distribution of the scale in DCE 2. We now explain how we tested these ideas statistically.

2.3. Brand measures and statistical model estimation

We used a sequential approach to estimate the effects of the DCE 1 brand-equity-associated brand utilities and the scale of utility in DCE 2. For DCE 1, we used the approach discussed in Louviere et al. (2008) to estimate individual-level-choice models using the most and least

---

3 We tested for differences due to priming by estimating a Multinomial Logit Model for the DCE 2 choices for each of the two data subsets from the different choice experiment sequences (i.e., DCE 1-DCE 2 and DCE 2-DCE 1). We also estimated a Multinomial Logit Model for the joint data and compared the differences in likelihoods obtained from the separate versus the joint estimation via a likelihood ratio test.

4 We use individual level models rather than a random coefficient approach, given that the latter imposes a structure across individuals. If this structure is not appropriate, it may lead to biased individual-level estimates (e.g., Fiebig et al., 2010), which could prove problematic for the test of interest in this paper.
preferred choices to develop a semi-ordering of the choice options in each choice set.

Louviere et al. (2008) describe how to convert the partial rankings that can be derived from the most and least preferred choice responses in our DCEs to the expected choice frequency counts that should be observed in all possible choice sets if a person chooses consistently with their ranking. For DCE 1, we calculated the expected choice counts from each individual’s implied ranking in each choice set given by their most and least preferred choices. The expected counts, equal “8” for the most preferred, “1” for the least preferred and “3” for the two middle options (“3” equals the average of the expected counts for the two middle rank orders; i.e., the average of “2” and “4”). We used the natural log of the expected counts as the dependent variable and the expected counts as the weights in a weighted least squares (WLS) regression for each person in DCE 1.

The statistical model used to estimate the effects of interest in DCE 2 is an extension of the G-MNL model proposed by Fiebig et al. (2010) and the RCGSMNL model of Swait and Erdem (2002). The G-MNL model allows for various forms of random scale heterogeneity across respondents or choice sets whereas the RCGSMNL captures the deterministic impact of exogenous variables on the utility scale of different alternatives in the same choice set. Our model jointly accounted for alternative- and individual-specific random variability in scale and preference heterogeneity, and allowed us to test our hypothesis of interest, namely, whether brand-equity-associated utility leads to more consistent choices in the product feature experiments. Like RCGSMNL, our model cannot be derived in analogy to the MNL model kernel by making conditional assumptions about the distribution of the error term. Regardless, it is a valid probabilistic choice model because its probabilities lie between zero and one, and sum to one over a choice set.

Table 1
Summary statistics of $R^2$ for individual-level WLS estimations of DCE 1.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airline</td>
<td>0.631</td>
<td>0.646</td>
<td>0.200</td>
<td>0.990</td>
</tr>
<tr>
<td>Car insurance</td>
<td>0.652</td>
<td>0.675</td>
<td>0.130</td>
<td>0.990</td>
</tr>
<tr>
<td>Listening devices</td>
<td>0.831</td>
<td>0.903</td>
<td>0.107</td>
<td>0.990</td>
</tr>
<tr>
<td>Visual entertainment devices</td>
<td>0.846</td>
<td>0.891</td>
<td>0.107</td>
<td>0.990</td>
</tr>
<tr>
<td>Auditory devices</td>
<td>0.853</td>
<td>0.913</td>
<td>0.087</td>
<td>0.990</td>
</tr>
<tr>
<td>After market add-on devices</td>
<td>0.845</td>
<td>0.907</td>
<td>0.087</td>
<td>0.990</td>
</tr>
</tbody>
</table>

Table 2
Parameter estimates from DCE 1, brand utility, and corresponding ranks for one respondent.

<table>
<thead>
<tr>
<th></th>
<th>Investment</th>
<th>Consistency</th>
<th>Credibility</th>
<th>Quality</th>
<th>Risk</th>
<th>Time</th>
<th>Brand utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qantas</td>
<td>0.10</td>
<td>0.45</td>
<td>0.20</td>
<td>0.14</td>
<td>0.10</td>
<td>0.04</td>
<td>1.03</td>
</tr>
<tr>
<td>Virgin Blue</td>
<td>-0.01</td>
<td>-0.01</td>
<td>0.00</td>
<td>0.08</td>
<td>0.06</td>
<td>-0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>JetStar</td>
<td>0.15</td>
<td>0.00</td>
<td>-0.08</td>
<td>0.08</td>
<td>0.10</td>
<td>0.08</td>
<td>0.33</td>
</tr>
<tr>
<td>Tiger</td>
<td>-0.25</td>
<td>-0.43</td>
<td>-0.11</td>
<td>-0.30</td>
<td>-0.26</td>
<td>-0.07</td>
<td>-1.43</td>
</tr>
</tbody>
</table>

Ranks

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Qantas</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Virgin Blue</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>JetStar</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Tiger</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

More specifically, we modeled the deterministic utility of person $n$ associated with alternative $j$ on purchase occasion (or choice scenario) $t$ as given by the following equation:

$$U_{njt} = \left[ \alpha_{njt} \beta + \gamma_{njt} + (1-\gamma)\sigma_{njt} \right] x_{njt}$$

where $x_{njt}$ is a vector of attributes of alternative $j$. This vector $x_{njt}$ includes all variables that affect the mean valuation of alternatives, namely, the attributes of choice options and alternative-specific constants. The alternative-specific constants captured possible systematic tendencies by some respondents to choose alternatives positioned on left or right-hand sides across all choice sets. $\beta$ is the vector of mean attribute utility weights in the population and $\gamma_{njt}$ is a person $n$ specific deviation from the mean, which captures preference heterogeneity. The parameter $\sigma_{njt}$ can be interpreted as the scale of the error term, which we allowed to vary across respondents, alternatives, and choice sets. The higher that this parameter is for an individual, the more consistent they are in the utility evaluation of a particular alternative.

The parameter $\gamma$ lies between 0 and 1 and governs how the variance of residual preference heterogeneity varies with scale in a model that includes both. For $\gamma$ close to 1, differences in scale across individuals do not lead to a wider spread of the vector of utility weights, and thus, the standard deviation of the residual preference heterogeneity $\eta_{njt}$ is independent of the scaling. In contrast, for $\gamma$ close to 0, the standard deviation of $\eta_{njt}$ is proportional to the scale $\sigma_{njt}$. Note that the model parameters are neither identified for $\gamma = 0$ nor for $\gamma = 1$, given that in these cases, the scale $\sigma_{njt}$ cannot be separated from the preferences $\beta$ and $\eta_{njt}$, or $\gamma$ cannot be separated from the individual deviations from the preferences $\eta_{njt}$, respectively.

We modeled taste heterogeneity by assuming a multivariate normal, $\text{MVN}(0, \Sigma)$, distribution for the $\eta$ vector with diagonal covariance matrix, $\Sigma$. The scale parameter must be restricted to be strictly positive, so we modeled scale heterogeneity by assuming a lognormal distribution (see also Fiebig et al., 2010). We completed the scale distribution specification by allowing $\sigma_{njt}$ to be affected by a vector $z_{njt}$ of attributes that influence scale, described as follows:

$$\sigma_{njt} = \exp(\sigma + \theta z_{njt} + \tau e_{njt})$$

where $e_{njt} \sim N(0,1)$.

For the test of interest in this paper, we used the brand utility rankings from DCE 1 to define the vector $z_{njt}$ (thus, $z_{njt}$ reduces to $z_{nj}$). If higher brand-equity-associated brand utility indeed results in
significantly higher scales (i.e., in more certain choices), the parameter θ should be significant and positive.

We specified the choice probabilities for alternative i on purchase occasion (or choice scenario) t for person n as follows:

\[
P_{ni} = \sum_{j} \exp \left( \alpha_{njt} + \gamma_{njt} + (1 - γ)\alpha_{njt} \right) \prod_{k} \phi(\varepsilon_{nk}(0.1)\theta(\gamma(0.1)\beta_{nk}(0, 1))\phi(\gamma(0, 0.1)\Sigma(0, 0.0)\delta_{nk}(0, 0.0)\phi(\Sigma(0, 0.0)\delta_{nk}(0, 0.0))) \right)
\]

(3)

At this point, it should be noted that our model nests the MIXL model (Revelt & Train, 1998) for \( \sigma_{njt} = 1 \). For the G-MNL model, the parameter \( \sigma_{njt} \) must be the same across alternatives (i.e., \( \sigma_{njt} = \sigma_{njt} \)), and \( \gamma = 0 \) and \( \tau = 0 \) yields the RCSSMNL model of Swait and Erdem (2002). Finally, our model also nests a model where differences in utility are only due to differences in scale for \( \tau_{njt} = 0 \).

We followed Fiebig et al. (2010) and set \( \alpha = -\tau^{2}/2\sigma_{njt} \) that the mean of \( \varepsilon_{njt} \) equals 1 for \( \theta = 0 \) for identification purposes. We estimated the parameters of Eqs. (1) and (2) using simulated maximum likelihood with log-likelihood function as follows:

\[
L = \prod_{t} \sum_{i} \ln \left( \int_{\theta} \left[ \prod_{j} \left[ \prod_{k} \phi(\varepsilon_{nk}(0.1)\theta(\gamma(0.1)\beta_{nk}(0, 1))\phi(\gamma(0, 0.1)\Sigma(0, 0.0)\delta_{nk}(0, 0.0)\phi(\Sigma(0, 0.0)\delta_{nk}(0, 0.0))) \right) \right) \right)
\]

(4)

where \( i \) refers to the chosen alternative at the \( t \)-th choice occasion (\( t = 1, ..., T \)).

3. Results

As outlined above, we tested our hypothesis that the brand-equity-associated brand utility (as measured in DCE 1) affects consumers’ choices in product feature experiments (i.e., DCE 2) by defining a model with the choice alternative’s attributes as elements of \( x_{njt} \) and the DCE 1 brand-equity-associated brand utility rankings as the only variables in \( z_{njt} \) (Model 1). We tested our model against the following three competing models: a model that does not include any covariates as explanatory variables for scale and thus equals the GMNL (Model 0), a model that includes the rankings of the brand equity underlying constructs (Model 2), and a model that includes brand dummies as covariates for scale (Model 3).

Model 0 served only for comparison purposes to test if scale differences across brands indeed play an important role in explaining choices in DCE 2. Model 2 allowed for the possibility that consumers differ in the weights that they attribute to different components of brand signals. In contrast, Model 3 tested whether differences in error variances (scale) across brands were constant across consumers but do not depend on information obtained in DCE 1.

Table 3 lists the estimated parameters for each model (these data are not available for Model 0) and the associated BIC. For the airline and the car insurance datasets where we have hold-out choice sets available, we also report the root mean squared error (RMSE) of the predicted choice probabilities versus the actual choice proportions averaged across all respondents.

The results in Table 3 supported our hypothesis that in all categories brand-equity-associated brand utility ranks (as measured in DCE 1) had a positive impact on the scale of choices in the product feature experiment (i.e., DCE 2). For Model 1, this impact was significant in all categories except airline choices. Thus, higher brand-equity-associated brand utility measures in DCE 1 were associated with higher utility scales in DCE 2, which is consistent with the brands acting as signals to reduce the need to review previously formed preferences (as suggested by Erdem & Swait, 1998). However, the model fits, as measured by BIC and by hold-out RMSE (where applicable), only improved compared with Model 0 in two categories, namely, visual entertainment devices and auditory devices.

The competing Model 2 includes the individual construct rankings instead of aggregate brand utility ranking as explanatory factors for scale, and we found that the more complex model only fits the data for the listening devices category better in terms of BIC than the model that only includes the brand utility ranking as a covariate for scale. Yet, even in this category, Model 2 still performed worse than the status quo model (Model 0) that included no covariates, implying that the additional complexity did not add explanatory power. However, Model 2 confirmed Swait and Erdem’s (2007) finding that a brand’s credibility leads to more consistent choices, that is, for four of the six categories, there was a significant and positive impact of a brand’s credibility rank on scale whereas in the remaining two categories, the impact of credibility on scale was not significant. Apart from a positive (and, in two categories, significant) impact of a brand’s investment on the scale of utility, we found that the remaining five constructs had no consistent impact on scale across categories.

Finally, we compared Model 1 to Model 3, which allowed for brand-specific scale effects not related to the brand utilities from DCE 1. We found that this model performed worse than Model 1 in terms of BIC and hold-out RMSE (where applicable) for all categories, except airline and listening devices. This result implies that the impact of a brand on respondents’ choice consistency is specific to an individual and depends on respondents’ perceived brand equity. It is worth noting, however, that Model 3 performed best in the airline category, suggesting that respondents’ choice consistency in this category was brand-specific; however, the information on perceived brand equity obtained in DCE1 was not necessary to explain consumers’ choices.

4. Discussion and conclusions

We studied the impact of Erdem and Swait’s (1998) brand-equity-associated brand utility measures on product choices. We combined two DCEs to test the association between a brand-anchored DCE (used to measure brand utility) and a brand-feature DCE (which varied key features of real brands). For five out of the six categories, we found a significant positive impact of the brand-equity-associated brand utility rankings on the scale of utility in the choice data. Thus, our results indicate that higher brand-equity-associated brand utility measures significantly increased respondents’ utility scale. Thus, stronger brands led to more consistent choices, supporting Erdem and Swait’s (1998) view of brands as information signals. In turn, this finding suggests that stronger brand signals can help simplify consumer decision processes and/or make consumers more confident about their choices, leading to more consistent choices. However, our results also showed that model fit only improved in two of the six categories, implying that the impact of brand equity on scale may have limited importance in explaining choice behavior. Similarly, given that we could not completely rule out priming effects that may arise from the brand equity experiment DCE 1, the extent to which brands affect scale may be lower in real-world choice settings. The last two issues should be the subject of further research to clarify and generalize this study’s findings.
Together with prior work on the impact of brands in consideration set formation (Erdem & Swait, 1998), our results suggest that stronger utility brands can help guide consumers’ purchase decisions. Retailer assortment sizes have greatly increased in many markets, possibly making choices more difficult for some consumers (Botti & Iyengar, 2006; Chernev, 2003; Huffman & Kahn, 1998). Our results imply that marketing managers not only can increase brand-equity-associated brand utilities in ways shown by previous research but also may be able to deal with choice complexity by developing and maintaining stronger utility brands and transmitting consistent and credible signals about their quality. Given that credibility, as well as the other underlying brand equity constructs, is influenced by managerial actions affecting the brand, one take-away from our results is that marketing actions by managers can also affect the choice consistency of a market. However, it is important to note that higher choice consistency—or “preference discrimination”, as Swait and Erdem (2007) refer to it—does not imply that consumers are more likely to purchase a stronger utility brand. Rather, it means that (all else equal), stronger brands will exhibit more extreme choice probabilities. Thus, consumers will be more likely to choose higher utility brands and will be less likely to choose lower utility brands. This finding implies that it will be easier for strong brands to retain consumers as “loyal” customers and harder for competitors to switch those consumers’ loyalties.

Our results also suggested that brand-equity-associated brand utility measures positively affect scale and that this finding was robust across several data sources. This result is consistent with the findings of Swait and Erdem (2002), who showed a positive impact of marketing mix consistency (theoretically linked to brand credibility) on scale for a frequently moving consumer good (FMCG) product. Thus, our results suggest that marketing managers should focus on building and maintaining strong brands that send consistent and credible product quality signals, which can build brand loyalty via increasing choice consistency. As Erdem and Swait (1998) noted, consumer brand equity exists regardless of a brand’s quality position; hence, low- or high-quality brands can have relatively high equity with a consistent low-/high-quality position. Thus, focusing on consistent and credible signals should benefit any brand.

The signaling perspective on equity also suggests that firms can build stronger utility brands by communicating a long-term commitment to them, again requiring consistency within and across marketing mix components. For example, consistency is likely to be particularly important for product line extension decisions that are implicated in SKU proliferation. That is, increased choice complexity may be associated with within-brand SKU proliferation because brands assumed to be strong from the firm’s perspective can overstretch their signals via product line extensions. Thus, our work suggests that managers should ensure the consistency of quality signals generated by new extensions matches those of their parent brands so as not to confuse consumers and dilute brand signal (see also Wernerfelt (1987) for how umbrella branding may serve as a signal of a new product’s quality). Prior research on “brand fits” when extending into new categories clearly suggest that consumers can become confused and not “see” the fit of some extensions unless great care is taken (Aaker & Keller, 1990, Völckner & Baltzer, 2017).

Our work also contributes to the ongoing discussion about positioning private labels in retailing. In particular, recent research suggests that “consumers use their experience with one private label brand to update their beliefs about rival retailers’ brands, and [that] these effects are quite sizeable” (Szynamowski & Gijsbrechts, 2012); hence, controlling consumer perceptions of signals from private labels...
should be more difficult as the number of categories with such labels increases. Even though Szymanowski and Gijsbrechts (2012) find that the consumption of one private label brand reduces consumers’ uncertainty of others, they note that this result may be due to the similar positioning of the private labels in their study. Indeed, such a “familiarity spillover” may lead to more consumer uncertainty if consumers compare experiences with different quality tiers of private label brands. Thus, the recommendation of Geyskens, Gielens, and Gijsbrechts (2010), which cautions against cannibalizing effects of private label introductions (with variable levels of quality) being perceived as similar, might also reduce the perception of noisy signals by consumers. That is, retailers may wish to delink different quality-tier private labels by positioning them in different shelf areas or creating stand-alone brands rather than sub-brands under a retailer’s name.

Our work also provides important insights for the management of online consumer ratings. More than 60% of the respondents in the 2007 Forrester Research online survey reported that they sought user ratings, and as a response, consumer-review websites and e-commerce websites, such as http://www.yelp.com/, http://www.walmart.com/ and http://www.amazon.com/, make the distribution of ratings available on their websites (Sun, 2012). From an information economics perspective, these ratings provide important signals to consumers about a product’s quality. Interestingly, heterogeneity in the ratings distribution does not necessarily mean that consumers receive noisier signals. Instead, Sun (2012) shows that such a high variance may enable consumers with niche preferences to find a better matching product. This finding suggests that, in the case of product ratings and reviews, signal noise has to be conceptualized at the preference segment level and that only signals within one segment should be consistent. Thus, marketing managers may find it beneficial to carefully monitor the variance of consumer reviews and to sometimes rely on reporting only averages rather than the full rating distribution.

Our work also suggests future research opportunities. First, one should test if the effect of brand-equity-associated brand utility measures on scale also holds for choices in real markets. For example, one can combine data from experiments such as DCE 1 with data on actual choices reported in surveys or scanner panels (see, e.g., Horsky et al. 2006). Moreover, combining experiments such as DCE 1 with longitudinal purchase data would permit one to test some of the issues related to how brand-equity-associated brand utilities affect brand loyalty. Similarly, it would also be interesting and important to track the evolution of brand-equity-associated brand utility measures and their associated underlying constructs over time to determine whether and how different marketing campaigns associated with the brands involved affect each construct, as well as total brand utility.

Finally, it would also be useful to determine in what ways and to what degree characteristics of consumers and product categories influence the impacts of brand signals. For example, the signaling process is less likely to be successful if a signal receiver is not looking for the signal (e.g., in markets where consumers have sufficient experience with the product offerings) or does not know what to look for (e.g., in completely new product categories) (see, e.g., Connelly, Certo, Ireland, & Reutzel, 2011). Research involving a broader range of categories may be able to determine these relationships.

### Appendix A

#### Appendix A.1. Screenshot of the brand anchored design in the airline category

| 1: how much each company invests in its offerings | like Qantas | like Virgin Blue | like Jetstar | like Tiger Airways |
| 2: how consistent the products & services of each company are | like Tiger Airways | like Jetstar | like Virgin Blue | like Qantas |
| 3: how credible each company is | like Tiger Airways | like Jetstar | like Virgin Blue | like Qantas |
| 4: how high the quality of a company’s products & service are | like Qantas | like Virgin Blue | like Jetstar | like Tiger Airways |
| 5: how risky it is to choose each company’s products & services | like Tiger Airways | like Jetstar | like Virgin Blue | like Qantas |
| 6: how much time & effort you can save by choosing each company | like Qantas | like Virgin Blue | like Jetstar | like Tiger Airways |

To review each factor again, please [click here](#).

1. The airline I most prefer is (check only one)
   - [ ] Airline A
   - [ ] Airline B
   - [ ] Airline C
   - [ ] Airline D

2. The airline I least prefer is (check only one)
   - [ ] Airline A
   - [ ] Airline B
   - [ ] Airline C
   - [ ] Airline D
Appendix A.2. Screenshot of the product feature experiment in the airline category

<table>
<thead>
<tr>
<th>Features</th>
<th>Package A</th>
<th>Package B</th>
<th>Package C</th>
<th>Package D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Airline</strong></td>
<td>Virgin Blue</td>
<td>Jetstar</td>
<td>Qantas</td>
<td>Tiger Airways</td>
</tr>
<tr>
<td><strong>Return economy airfare</strong></td>
<td>$450</td>
<td>$550</td>
<td>$750</td>
<td>$650</td>
</tr>
<tr>
<td><strong>Flying time</strong></td>
<td>5 hours</td>
<td>7 hours</td>
<td>7 hours</td>
<td>5 hours</td>
</tr>
<tr>
<td><strong>Ticket change free of charge</strong></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Free in-flight food/beverages</strong></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Free In-flight alcohol</strong></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Stops</strong></td>
<td>none</td>
<td>1 stop</td>
<td>1 stop</td>
<td>none</td>
</tr>
</tbody>
</table>

1. Which trip package would you be most likely to choose?
   - Package A
   - Package B
   - Package C
   - Package D

2. Which trip package would you be least likely to choose?
   - Package A
   - Package B
   - Package C
   - Package D

3. Would you actually choose any of the trip packages described if they were available when you book your next cross-country trip in Australia?
   - Yes
   - No

Appendix B

The 7 constructs were described as follows:

1) how much each option invests in its offerings:
   - The company is constantly evolving
   - The company makes sure it keeps up to date
   - The company is a leader in using technology

2) how consistent the products and services of each option are:
   - You always get what you expect from the company
   - The price is always in line with what you'd expect from the company
   - The company's prices match to its overall image

3) how clearly each option communicates its products and services:
   - The company sends clear messages about what they offer
   - The company sends clear messages about their image
   - The company provides information that tells me clearly what to expect

4) how credible each option is:
   - The company is upfront/open/honest about its capabilities
   - The company has a name you can trust
   - The company has a good reputation

5) how high the quality of option's products and service are:
   - You can depend on the company's services
   - I feel confident that the company will serve me well
   - In terms of overall quality, I'd rate the company highly

6) how risky it is to choose each option's products and services:
   - I feel that whatever I bought from the company, it would perform well
   - I'm sure about the company
   - I know enough to feel comfortable with using the company

7) how much time and effort you can save by choosing each option:
   - You can quickly find out about the company
   - The company's website is easy to use
   - The company's website makes it quick to find information

References


Advanced brand concept maps: A new approach for evaluating the favorability of brand association networks

Oliver Schnittka a,1, Henrik Sattler a,*, Sebastian Zenker b,2

a Institute of Marketing and Media, University of Hamburg, Welckerstraße 8, D-20354 Hamburg, Germany
b Erasmus School of Economics, Erasmus University Rotterdam, P.O. Box 1738, Rotterdam, The Netherlands

Abstract

John, Loken, Kim, and Monga (2006) have introduced brand concept maps (BCM) as a powerful approach to measuring brand image according to the structure of the underlying brand association networks and reveal the strength and uniqueness of brand associations. Interestingly, BCM, as well as other consumer mapping techniques, do not incorporate explicit measures for the favorability of brand associations. This study extends the original BCM approach with explicit information on the favorability of single brand associations and, further, develops a new metric, brand association network value (BANV), which quantifies overall network favorability. Our advanced BCM approach and the new BANV metric are managerially relevant in that they allow for comparison of the favorability of networks at both individual brand association and aggregate network levels. We illustrate the relevance of our BANV metric within an empirical application and demonstrate its validity.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Brand image constitutes an important element of customer-based brand equity (Keller, 1993). Understanding brand image demands the identification of a network of strong, unique, and favorable brand associations because consumers store brand information in the form of associative networks (Anderson, 1983; John, Loken, Kim, & Monga, 2006; Keller, 1993). Brand association networks identify, for instance, which associations are directly or indirectly linked to the brand and how these brand associations are connected to one another. Association networks also indicate the brand’s value to consumers and suggest ways to leverage its equity in the marketplace (Aaker, 1996; John et al., 2006).

Two categories of techniques are specifically designed to measure brand association networks: consumer mapping techniques and analytical techniques (John et al., 2006). The former, including brand concept maps (BCM) and Zaltman’s metaphor elicitation technique (ZMET), elicit individual brand association networks directly from consumers (John et al., 2006; Zaltman & Coulter, 1995). That is, respondents reveal how their brand associations relate to the brand and others by constructing their own network of associations. With these individual maps, researchers can aggregate the information to produce a consensus brand association network. Analytical techniques instead uncover brand associations through consumer surveys (e.g., repertory grids) and employ analytical methods to reveal the underlying consensus brand association network (Henderson, Iacobucci, & Calder, 1998).

Among consumer mapping approaches, the BCM method is particularly promising (John et al., 2006). Unlike analytical techniques, such as network analysis (Joiner, 1998; Lynch & Srull, 1982), the BCM technique allows for the analysis of brand association networks at both individual and aggregate levels because brand maps emerge for each respondent. In contrast to other mapping techniques, such as ZMET, BCM also gathers consumer perceptions using structured association elicitation, mapping, and aggregation procedures, which result in an easy-to-administer, less costly, and less subjective approach (John et al., 2006). John et al. (2006) also offer empirical evidence of the high reliability and validity of the BCM approach.

Because brand image is defined by the strength, uniqueness, and favorability of brand associations, organized in a network (Keller, 1993), it is surprising that BCM and other techniques do not incorporate explicit measures of the favorability of brand associations. That is, the BCM technique identifies relevant brand associations, groups them in a network, and offers information on the uniqueness of the associations (e.g., in terms of the multitude of brand-specific associations in a brand map thereby assuming that additional associations in the brand map increase the probability that associations are unique in comparison with competitors) and their strength (e.g., in terms of the degree to which associations are directly linked to the brand node or the strength of the associations’ linkage to the brand node (e.g., weak, moderate, or strong)) but does not provide explicit favorability information.
Prior research reveals that the favorability of brand associations, as evaluated by consumers, varies substantially, particularly with regard to (1) individual evaluative judgments (i.e., how favorable each association is pronounced to be for the specific brand) and (2) their importance to the overall purchase decision (Fishbein & Ajzen, 1975; Keller, 1993; Wilkie & Pessemier, 1973). Therefore, we extend the original BCM approach by integrating explicit information on the favorability of brand associations within brand association networks (i.e., advanced BCM). Specifically, we include the original information about uniqueness and strength but also integrate explicit favorability information regarding (1) evaluative judgments of each brand association (Keller, 1993), as well as (2) the individual importance of each brand association to a consumer's purchase situation. This paper demonstrates that the added information has valuable management implications in that it makes the resulting networks more meaningful.

Because the original BCM technique and other mapping and analytical techniques do not account for explicit measures for the favorability of brand associations, they do not provide information regarding the favorability of the underlying association network. Therefore, we introduce a new metric, the brand association network value (BANV) metric, which, for the first time, quantifies the overall favorability of brand association networks by combining network structure (i.e., the uniqueness and strength of brand associations) and the favorability of single brand associations (i.e., the evaluative judgment and the importance to the specific purchase decision) into a single measure. The new metric enhances the usefulness of the BCM methodology for comparisons of the network favorability at individual and aggregate network levels.

In the next section, we briefly outline the original BCM approach and our extended advanced BCM approach and derive our new BANV metric. We then describe the research design for our empirical application and demonstrate the validity of our new metric. In the final section, we discuss how the advanced BCM technique and the BANV approach contribute to brand image measurement and derive further research applications and limitations.

2. Derivation of the advanced BCM approach and the brand association network value (BANV)

2.1. The original BCM procedure

To demonstrate the contrast of the advanced BCM approach to the original BCM procedure, introduced by John et al. (2006), we begin with a brief description of the latter. The original BCM procedure consists of two major stages that provide individual brand maps: During the elicitation stage, highly relevant brand associations are identified (e.g., through in-depth interviews). During the mapping stage, respondents are asked to develop an individual brand map with these predetermined brand associations, in the center of which the brand emblem appears. Respondents then assign different strengths of association links to connect the selected brand associations directly to the brand emblem or indirectly to one another.

Thus, the original BCM approach provides individual network information regarding (a) the presence of each of the predetermined brand associations on the brand map, (b) the level at which each association appears on the brand map (e.g., first-order association—directly connected to the brand, second-order association—connected under a first-order association), and (c) the strength of linkage connecting each association to the brand or to another association (i.e., weak links=single lines, moderate links=double lines, and strong links=triple lines).

Imagine the case of a premier health care brand, the Mayo Clinic, which is interested in using the original BCM approach to discover how patients perceive the brand after clinic stays (see John et al., 2006). Fig. 1a represents the original consensus map aggregated for 90 Mayo Clinic patients. Specifically, the solid-line circles signify core associations (i.e., associations that are included on at least 50% of the individual maps) and the dashed-line circles signify non-core associations (i.e., associations that are included on less than 50% of the individual maps).

The consensus map in Fig. 1a provides information about the structure of the network, including the level of single brand associations within the network (e.g., six first-order associations) and the strength of their connection to other brand associations or the brand itself (e.g., strong triple-line linkage between the Mayo Clinic and “Best doctors in the world”). Although this network information is relevant to managers in determining consumers’ brand perceptions, it does not explicitly indicate, for instance, which of the six first-order brand associations should be a priority in future marketing activities because no systematic information regarding the evaluative judgment of each association and the corresponding association’s importance in a choice decision for a specific clinic is available. Note that some associations, such as “Leader in medical research”, suggest relevance to a certain degree and thereby provide implicit favorability information, but the favorability of brand associations becomes not explicit for all relevant associations within the original BCM approach.

2.2. Extension of the original BCM approach (advanced BCM)

Because the original BCM approach does not provide any explicit favorability information regarding single brand associations, our advanced-BCM procedure includes an additional evaluation stage as a part of the mapping procedure regarding (1) consumers’ evaluative judgments of each brand association and (2) the individual importance of each brand association in the context of a consumer’s purchase situation.

The first extension comprises respondents’ statements regarding how favorable each association placed on an individual brand map is for the specific brand measured by a one-item, seven-point Likert scale (e.g., “The patient care available at the Mayo Clinic is good”). In contrast to the original BCM approach, in which some associations may have valences (e.g., “Best doctors in the world”), each association within our advanced BCM approach is neutrally titled and combined with the additional phrase “is good”. This process allows for positive and negative evaluations (i.e., 1=totally disagree; 7=totally agree).

Regarding the second extension, Keller (1993) argues that not all brand associations are relevant to the same extent for a purchase decision and, therefore, vary in their importance. For instance, a patient might strongly associate the Mayo Clinic with the association “World leader in new medical treatments” but rate the importance of this association as low in comparison with other associations, such as “Best patient care available” because s/he prefers a comfortable stay at the clinic. Thus, our advanced BCM also includes respondents’ statements regarding the individual importance of each brand association to a purchase (or choice) decision as reported on a one-item, seven-point Likert scale (e.g., “The patient care available is an important attribute when choosing a clinic”). Implicitly, the original BCM approach might include information on the importance of brand associations. For instance, associations having a high degree of network centrality (e.g., being more interconnected) might be interpreted as being more important for understanding a consumer’s brand image. However, the
Consensus maps: Mayo Clinic (see John et al., 2006, p. 556)

a) Original-BCM (N= 90)

b) Advanced-BCM (hypothetical)

Notes: The solid circles represent core associations of Mayo Clinic’s brand image, dashed-line circles represent non-core associations, and the single, double, and triple lines represent the average strength of the association linkage across all respondents, according the original BCM approach. For the advanced BCM approach, the diameter of each circle represents the average individual importance of an association for the overall clinic evaluation within a clinic choice situation, whereby a larger diameter indicates greater importance. The darkness of each circle represents the average evaluative judgment of an association, with darker circles indicating a more favorable evaluation.

Fig. 1. Consensus maps: Mayo Clinic (see John et al., 2006, p. 556). a) Original-BCM (N=90). b) Advanced-BCM (hypothetical).

original BCM approach does not impart importance ratings regarding how consumers weight the association during a purchase decision.

For managers, such explicit information about the favorability of brand associations is important, particularly when they determine the extent to which they should focus on specific associations in their marketing activities. The advanced BCM procedure is hypothetically illustrated for the Mayo Clinic case in Fig. 1b. The diameter of each circle represents the average individual importance of an association for the overall product evaluation in a purchase situation (with a larger diameter indicating greater importance), and the darkness of each circle represents the average evaluative judgment of an association (with darker circles indicating a more favorable evaluation). Regarding the six first-order brand associations, the association “Best doctors in the world” has more importance for a decision to choose a clinic than the association “Leader in medical research.” Thus, the minor relevance of the latter association for all brands within the product category implies that marketing efforts should focus primarily on the doctors’ competence rather than on the clinic’s research and development competence because single brands are generally not expected to actively influence an association’s relevance within a product category (e.g., actively increasing the relevance of the association “Leader in medical research”). The figure also reveals that the association “Best patient care available” prompts less favorable evaluations than “Expert in treating serious illnesses.” Because single brands are expected to actively influence the evaluative judgment of their own brand associations through marketing communication, the Mayo Clinic might wish to primarily address their concern about patients’ well-being in its marketing communication to enhance consumers’ evaluations of the patient care association.

2.3. The brand association network value (BANV)

In the following section, we derive the brand association network value (BANV), which, for the first time, quantifies the overall favorability of brand association networks. The BANV is based on a widely accepted multi-attribute attitude model approach developed by Fishbein and Ajzen (1975), which suggests that brand attitudes are a multiplicative function of the salient (cognitive) beliefs that a consumer has about the product or service (i.e., the extent or probability to which consumers think that the brand is linked to certain attributes or benefits) and the evaluative (affective) judgment of those beliefs (Ajzen &
Fishbein, 1980; Fishbein & Ajzen, 1975; Wilkie & Pessemier, 1973). Therefore, in the context of BCM, we propose that the overall evaluation of the favorability of an individual brand association network for consumer $j$ (i.e., BANV$_j$) initially consists of the added evaluative judgments of all brand associations $m$ that appear in the individual brand map, multiplied with the corresponding strength of the association linkage to the brand node or superordinate co-associations in consumer memory (i.e., $\sum_{a=1}^{m} E_{aj} \cdot S_{aj}$; Reder & Anderson, 1980; John et al., 2006).

$E_{aj}$ represents the evaluative judgment of each brand association $a$ stated by respondent $j$ on a one-item, seven-point Likert scale, with higher scores indicating a more favorable rating. Furthermore, $S_{aj}$ represents the strength of an association’s direct linkage (i.e., weak, moderate, or strong) to its superordinate associations within the network or the brand node, with stronger linkages producing a greater association impact on the BANV.

Specifically, the activation of one information node (i.e., brand association) in one’s memory is generally expected to spread to other information nodes that are structurally linked (Collins & Loftus, 1975; Paivio, 1969). In turn, stronger links between individual information nodes increase the likelihood of the spread of activation between the nodes (Keller, 2008). Because superordinate information nodes serve as evaluation anchors for subordinate information nodes (Esch, Schmitt, Redler, & Langner, 2009; Iversky & Kahneman, 1974), a stronger linkage between an association and superordinate information nodes (i.e., superordinate associations or the brand node) will increase the likelihood of the activation of that association when the superordinate information nodes are stimulated. Therefore, according to Fishbein and Ajzen (1975), the evaluation of each association (i.e., $E_{aj}$) can be weighted by the strength of its direct linkage $S_{aj}$ (i.e., weak, moderate, or strong) to its superordinate associations within the network or the brand node.

To match the measure of $E_{aj}$, $S_{aj}$ is transformed to a seven-point scale, for which the best $S_{aj}$ values (i.e., a strong association link to the brand node or superordinate co-association directly connected with the brand) are coded as 7, the worst values are coded as 1, and all other values are interpolated accordingly.

According to Keller (1993), not all brand associations are relevant to the same extent to a purchase and, accordingly, a choice decision. Therefore, we weight evaluations of the favorability of each brand association, $E_{aj}$, according to the individual importance of that association to a purchase decision, $I_{aj}$, measured on a one-item, seven-point Likert scale, with higher scores indicating a more favorable rating. Although prior studies discussed the usefulness of simultaneously including information regarding evaluative judgments and the importance of brand associations in multi-attribute attitude models (see Bass & Wilkie, 1973; Wilkie & Pessemier, 1973), other studies, such as Hansen (1969), Lehmann (1971), or Bass, Pessemier, and Lehmann (1972) empirically revealed that the inclusion of importance weights actually increased the model’s predictive power. Building on the latter empirical findings and the conceptual framework conducted by Keller (1993), we expect evaluative judgments and importance to be meaningfully different. In addition, because predictive power can be expected to increase if the importance weights are elicited in a context of a specific purchase situation (rather than for a product class in general; Cohen & Ahtola, 1971; Wilkie & Pessemier, 1973), we decided to integrate such purchase situation-specific importance weights.

Considering that brand associations are organized in a network (Keller, 1993), associations may vary not only in their importance to consumers’ overall evaluation of a product but also in the lengths of their associative pathways to the brand node and the extent to which they are directly associated with the brand node (e.g., direct, first-order linkage of the attribute “Best doctors in the world” and indirect, second-order linkage of the attribute “Treats famous people” with the Mayo Clinic in Fig. 1a; John et al., 2006). Information processing theory assumes the existence of hierarchical storage, indicating that relevant information is linked more directly to the brand and can be retrieved more easily than subordinate information (Bettmann, 1979; Miller, 1956). In the context of BCM, we expect that an association that is closer to the brand node and, therefore, on a higher association level, $L_{aj}$, on the individual brand map, leads to a stronger association impact on the overall brand image.

Thus, we also weight the evaluation of each brand association, $E_{aj}$, according to the level of its placement, $L_{aj}$, within the network and transformed $L_{aj}$ to a seven-point scale to match the measures of $E_{aj}$, $S_{aj}$, and $I_{aj}$. Specifically, the best $L_{aj}$ values (i.e., association directly connected to the brand) are coded as 7, the worst values are coded as 1, and all other values are interpolated accordingly.

Finally, the BCM method recognizes brand-specific associations within a network that generate a sustainable competitive advantage, as well as a unique selling proposition (Broniarczyk & Alba, 1994; Keller, 1993; Ries & Trout, 1979). Thus, the multitude of brand-specific associations within a network $m$ might provide a proxy for the level of uniqueness of brand associations within a network. Because Krishnan (1996) empirically reveals that high- (low-) equity brands are characterized by a greater (smaller) number of associations on the respective brand maps, we predict that the multitude of brand-specific associations will have a positive impact on the overall BANV score.

Overall, our new BANV measure is composed of different dimensions of information that reveal the favorability (i.e., $E_{aj}$ and $L_{aj}$), strength (i.e., $I_{aj}$ and $S_{aj}$), and uniqueness (i.e., $m$) of brand associations, according to the following equation:

$$\text{BANV}_j = \sum_{a=1}^{m} E_{aj} \cdot S_{aj} \cdot I_{aj} \cdot L_{aj}$$

(1)

3. Research design

3.1. Procedure

We undertook two lab experiments at a German university, using a final sample of 111 respondents (study 1: average age: 24.7 years, 53.2% female) for the first study and 123 respondents (study 2: average age: 18.2 years, 56.9% female) for the second study. For our first study, cars provide the product category and the Volkswagen Golf the brand. To participate in the survey, respondents needed to have a minimum level of involvement with cars in the compact class segment. The Volkswagen Golf is the most familiar product brand within this segment, with uniformly high familiarity scores throughout the sample ($M_{\text{familiarity}}=4.54$, $SD=1.74$), which aids in generating a meaningful BCM. Furthermore, we used the Volkswagen Golf product brand instead of the corporate brand Volkswagen because heterogeneous perceptions are more likely to exist for a corporate brand than for a product brand. For our second study, sports shoes provide the product category and Adidas the brand. Adidas is the most familiar brand within this segment, having high familiarity scores in the sample ($M_{\text{familiarity}}=4.43$, $SD=1.29$).

The BCM approach (John et al., 2006) in both studies consists of two survey stages to provide individual brand maps, elicitation and mapping. During the elicitation stage of study 1, we identified highly relevant brand associations through in-depth interviews with 35 respondents who were very familiar with the compact car class. This pretest generated 121 relevant brand associations, from which we selected the top 25 to create a relevant association set for the mapping procedure. We eliminated any associations that could appear redundant. Thus, we developed the following list of associations: comfort, commodiousness, design, driving pleasure, durability, environmental sustainability, fuel consumption, being down-to-earth, handling, innovation, mobility, practicality, prestige, price, price–performance ratio, prominence, quality, reliability, resale value, safety, satisfaction, emotional commitment, trustworthiness, usualness, and youthfulness.
Similarly, during the elicitation stage of study 2, we identified highly relevant brand associations through in-depth interviews with Adidas marketing managers who were very familiar with the sports shoes industry. We again selected the top 25 associations to create a relevant association set for the mapping procedure: activeness, authenticity, comfort, creativity, credibility, encouragement, excitement, feel, friendliness, fun, individuality, innovation, inspiration, internationality, visibility, motivation, passion, provision of recognition, quality, strong-mindedness, style, support, trend setting, people uniting, vitality, and top athlete use.

In both studies, we then initiated the mapping stage of the BCM procedure, as described by John et al. (2006), by asking respondents to develop an individual brand map in which the Volkswagen Golf (Adidas) emblem appeared in the center. In addition to the 25 predetermined brand associations, respondents received five blank cards, which they could use to write down additional relevant associations. Finally, they assigned different association links (i.e., weak links = single lines, moderate links = double lines, and strong links = triple lines) to connect the selected associations and their own brand associations with the Volkswagen Golf (Adidas) and with one another.

Moreover, we added an evaluation stage. Respondents indicated their evaluative judgment, $E_{ij}$, of each association within the brand association network for the Volkswagen Golf (Adidas) on a one-item, seven-point Likert scale (e.g., “The fuel consumption of the Volkswagen Golf is good” or “The comfort of an Adidas sports shoe is good”), whereby a higher score indicated a more favorable rating. They also indicated individual importance, $I_{ij}$, for the overall product evaluation for a car rental (sports shoes purchase) situation, again with a one-item, seven-point Likert scale (e.g., “Fuel consumption is an important product attribute when renting a car” or “Comfort is an important attribute when purchasing sports shoes”).

### 3.2. Scales

To assess the nominal validity and predictive validity of the BANV, we measured respondents’ attitudes toward the brand, brand familiarity, and their purchase intention for the brand using several seven-point Likert scales on which higher levels indicated more favorable ratings. To measure respondents’ attitudes toward Volkswagen Golf (Adidas), we used three items from Osgood, Suci, and Tannenbaum (1957): good, positive, and favorable ($\alpha = .90$). The measure of familiarity consisted of three items adapted from Kent and Allen (1994): familiar, experienced, and knowledgeable ($\alpha = .92$). We also measured respondents’ purchase intentions with regard to the Volkswagen Golf (Adidas) on a three-item scale, adapted from Dodds, Monroe, and Grewal (1991), that included consider buying, likelihood to purchase, and willingness to buy ($\alpha = .83$). Thus, all three validity measures achieved good reliability (Cronbach’s alpha) in both studies ($N=234$ respondents). Furthermore, we used three seven-point Likert scale items adapted from Mittal (1995) as a measure of involvement: important, concerned, and care about ($\alpha = .95$).

To test the predictive validity of the BANV metric regarding consumers’ choice probability for the Volkswagen Golf, we employed a choice-based conjoint analysis in study 1, with name and price as the presumably most important attributes. To create a simulated choice situation that was as realistic as possible, we implemented a choice scenario in which respondents wanted to rent a car for a weekend (instead of purchasing a new car). The brand attribute levels cited four well-known competitors in the German compact class market: the Volkswagen Golf, the Audi A3, the BMW One, and the Mercedes A-Class. For price, we selected four common market prices for renting a car for a weekend: €79.99, €94.99, €109.99, and €124.99. We used eight choice sets and two holdout sets, each with three alternatives, in a design that accounts for the efficiency criteria of balance, orthogonality, and minimal overlap (Huber & Zwerina, 1996). A reliability test using the two holdout sets revealed inconsistent response behavior for ten participants, whom we excluded from further analysis. We applied a multinomial logit (MNL) model to estimate the consumer choice probabilities. The MNL models reveal the preferences $\beta$ for each attribute level in terms of choice probabilities $p$, whereby an alternative i chosen from a set of J alternatives would be integrated into a likelihood function as follows:

$$ p(i | J) = \frac{\exp(\beta x_i)}{\sum_{j=1}^{J} \exp(\beta x_j)} \tag{2} $$

where the vector $x_i$ describes the characteristics of alternative $j$ (i.e., specific levels of attributes). To obtain $\beta$, we applied a hierarchical Bayes estimation that yields individual-level estimates (Rossi & Allenby, 1993).

To measure consumers’ choice probability for Adidas, we employed a direct measure of brand choice in study 2. Specifically, we embedded a simulated choice situation for sports shoes in which consumers had to decide between Adidas and three well-known competitors: Nike, Puma, and Reebok (and assigned a choice probability of 1 if the Adidas sports shoe was chosen and 0 otherwise).

### 4. Empirical results

#### 4.1. Empirical application of the advanced BCM approach

Our empirical implementation of the advanced BCM approach to the Volkswagen Golf case produces the consensus map in Fig. 2b (the corresponding consensus maps for the Adidas case can be provided on request). To illustrate the additional information value of the advanced BCM approach in comparison with the original BCM approach, we derive a corresponding BCM consensus map in Fig. 2a using the original approach (John et al., 2006).

Both approaches indicate the same six first-order associations and four second-order associations, as well as the same average strength of linkages between the associations. The advanced BCM approach also contains information regarding the importance and evaluative judgments of associations, as is evident in Fig. 2b. This additional information is of high managerial relevance. For example, the original BCM in Fig. 2a suggests that the association “prominence” should be strongly addressed in future marketing activities because it shows the strongest (i.e., triple-line) direct linkage with the Volkswagen Golf brand and should contribute to building and leveraging brand image. However, the advanced BCM consensus map in Fig. 2b reveals that the “prominence” association has only minor importance for purchase decisions in comparison with the other first-order associations. Consequently, this association should not be a primary topic in future marketing activities, especially because single brands are not expected to actively increase associations’ relevance within a product category. Instead, managers should focus on the first-order associations “price-
performance ratio” and “safety,” which have much greater relevance for consumers thinking of purchasing a car.

The original BCM approach would also imply that “reliability,” “safety,” “being down-to-earth,” and “price–performance ratio” contribute equally to the Volkswagen Golf’s brand image in that they all indicate moderate (i.e., double-line), direct association linkages with the brand node. This map provides no evidence regarding which association should be primarily targeted in future marketing activities. In contrast, the advanced BCM consensus map in Fig. 2b reveals that “price–performance ratio” provokes less favorable evaluations than the remaining first-order associations. Furthermore, the second-order association “price” is evaluated even less favorably (i.e., too expensive), which implies that the unfavorable price–performance ratio evaluation is caused by an overly expensive price for the Volkswagen Golf. This finding suggests that Volkswagen should review its pricing for the Golf because its current level not only deteriorates consumers’ evaluation of the second-order price association but also spills over to the first-order price–performance ratio association. Such important implications for the strategic management of specific brand associations cannot be elicited with the original BCM approach.

**Fig. 2.** Consensus maps: Volkswagen Golf. a) Original-BCM ($N=111$). b) Advanced-BCM ($N=111$).
4.2. Empirical application and validation of BANV

We transformed \( L_{aj} \) and \( S_{aj} \) to seven-point scales to match the measures of \( E_{aj} \) and \( I_{aj} \). The best values of \( L_{aj} \) (i.e., association directly connected with the brand) and \( S_{aj} \) (i.e., strong association link to the brand node or superordinate co-association) were coded as 7, and the worst values took on the value of 1, with all other values interpolated accordingly. We thus assume a linear progression for \( L_{aj} \) and \( S_{aj} \) (John et al., 2006). Furthermore, we assume that associations of subordinate order (e.g., second-order) are weighted by the strength of their linkage \( (S_{aj}) \) to superordinate associations (e.g., first-order). If an association is linked to more than one superordinate association, we used the average value of their linkages.

In addition, we analyze the intercorrelations between the four BANV components \( E_{aj}, I_{aj}, L_{aj}, \) and \( S_{aj} \) for all brand associations (\( N=1609 \)), which are included on the 234 individual brand maps for both studies to ensure empirical distinctiveness among themselves. All correlation coefficients are smaller than .25, which indicates an at least satisfying level of discriminant validity (see Table 1).

We first analyze the plausibility of our BANV metric in terms of face validity to determine if the observed results are intuitively plausible and whether the visual operationalization of the brand maps offers an adequate translation of the BANV construct. Krishnan (1996) demonstrates empirically that high-equity brands are characterized by more associations in their brand maps. John et al. (2006) indicate that highly familiar brands (which should also have more favorable brand images, according to the mere exposure effect; Anand, Holbrook, & Stephens, 1988; Zajonc, 1968) produce brand association networks with more brand associations in general, more first-order brand associations and brand association links, as well as stronger brand association links. Our empirical results suggest that our BANV metric has good face validity in both studies because it positively correlates with the number of brand associations in general \((r=.84, p<.00, N=234)\), the number of first-order brand associations \((r=.52, p<.00, N=234)\), and the number of brand association links in general \((r=.42, p<.00, N=234)\) as well as the number of strong brand association links \((r=.41, p<.00, N=234)\).

To test the nomological validity of the BANV metric, we rely on Keller’s (1993) assertion that consumers’ overall attitudes toward a brand represents an essential component of their overall brand images. Consumers with more favorable overall attitudes toward a brand should thus have a more favorable BANV score. Our results consistently reveal a positive correlation between our BANV metric and consumers’ overall attitude toward the brand for both studies \((r=.40, p<.00, N=234)\).

According to mere exposure theory, repeated stimulus exposure improves consumers’ brand image, so their brand image should be enhanced by their brand familiarity. Because consumers who are highly familiar with a brand should have a more favorable brand image (Anand et al., 1988; Zajonc, 1968), we predict that they should exhibit more favorable BANV scores. In line with this expectation, the BANV metric reveals a positive correlation with consumers’ brand familiarity for both studies \((r=.36, p<.00, N=234)\).

Regarding predictive validity, in line with the theory of planned behavior (Ajzen, 1991), we expect that consumers with a more favorable BANV score express higher purchase intentions and higher choice probabilities for the brand in a simulated choice situation. Our results consistently reveal a positive correlation between the BANV metric and (1) consumers’ purchase intentions \((r=.35, p<.00, N=234)\) and (2) consumers’ choice probabilities within the conjoint procedure in favor of the investigated brand \((r=.24, p<.00, N=234)\) for both studies (the results of the pooled sample are in line with those of the Volkswagen Golf sample (study 1) and the Adidas sample (study 2)).

In addition, our BANV metric, which incorporates different pieces of information pertaining to the favorability (i.e., \( E_{aj} \) and \( I_{aj} \)), strength \((L_{aj} \) and \( S_{aj} \)), and uniqueness of brand associations (i.e., \( m \)) reveals higher levels of nomological validity and predictive validity in comparison with simpler metrics, which only contain some of the information (see Table 2).

More specifically, we consider BANV{E,S} (i.e., \( \sum_{a=1}^{m} E_{aj} \times S_{aj} \)), as an appropriate baseline model, according to Fishbein and Ajzen (1975); BANV{E,S} illustrates the evaluative judgment of each association, weighted by the strength of its linkage to the brand node or superordinate co-associations for all brand associations \( m \) that appear on the individual brand map. Compared to the baseline model BANV{E,S}, our results for both studies demonstrate that the inclusion of the individual importance of an association for a purchase decision \( L_{aj} \) in the model (i.e., \( \text{BANV}_m = \sum_{a=1}^{m} E_{aj} \times S_{aj} \times L_{aj} \)) leads to significantly higher correlations regarding the four validation criteria brand attitude \((z=5.49, p<.01, N=234)\), brand familiarity \((z=2.29, p<.05, N=234)\), and consumers’ purchase intentions \((z=3.67, p<.01, N=234)\) as well as choice probabilities \((z=2.68, p<.01, N=234)\) for each case, we used Steiger’s z-test-procedure for correlated correlation coefficients to estimate the differences between the respective correlation coefficients; Steiger, 1980). Thus, the inclusion of \( L_{aj} \) increases the power of the model in terms of nomological and predictive validity.

Furthermore, in comparison with BANV{E,S,I}, the additional inclusion of the level of an association’s placement, \( L_{aj} \), within the network (i.e., \( \text{BANV}_m = \sum_{a=1}^{m} E_{aj} \times S_{aj} \times I_{aj} \times L_{aj} \)) further improves validity regarding brand attitude and brand familiarity (i.e., for both studies, we found significantly higher correlations for BANV than for BANV{E,S,I} regarding brand attitude \((z=2.46, p<.05, N=234)\) and brand familiarity \((z=2.02, p<.05, N=234)\)). The same pattern of results occurs regarding purchase intentions and choice probabilities, but the differences between the correlations of BANV and BANV{E,S,I} are not statistically significant \((z=.81, p>.05, N=234)\) for both studies \((z=.78, p>.05, N=234)\).

Overall, our empirical results suggest high levels of face validity, nomological validity, and predictive validity for our proposed metric. In addition, the findings reveal that all pieces of information (i.e., \( E_{aj}, I_{aj}, L_{aj}, S_{aj} \) and \( m \)) are actually needed to sufficiently quantify the overall favorability of brand association networks because BANV led to higher levels of validity than simpler metrics. Moreover, the BANV metric provides additional managerial insights in comparison with existing consumer mapping procedures and analytical techniques. In particular, the BANV metric allows for comparisons of brand association networks across consumer segments or developments over time. For example, consider Fig. 3a and b regarding the Volkswagen Golf case. The two maps represent two different consumer segments with low (Fig. 3a) and high (Fig. 3b) involvement in cars \((M_{low_{involvement}}=1.54, SD=.57, N=56; M_{high_{involvement}}=4.38, SD=1.12, N=55; t=16.90, p<.01)\). Suppose that Volkswagen wonders which of the two segments they should target with prospective advertising campaigns. To support this decision, Volkswagen needs an overall evaluation of the two segments’ brand association networks.

The original BCM approach describes the two consumer segments in terms of the level of single brand associations or the strength of specific brand association linkages within the network. Such information does not provide additional insights because it only includes the strength and favorability of single brand associations for both studies.
not provide sufficient evidence to support Volkswagen’s advertising campaign decision because the brand maps in Fig. 3a and 3b include the same number of core brand associations and first-order associations, as well as approximately the same number of weak, moderate, and strong brand association links. Thus, whether the brand image of the two segments differs substantially at the aggregate level in terms of their brand association network value remains unclear. However, the application of the BANV metric reveals that the overall network evaluation of highly involved consumers is significantly higher than that of less involved consumers (MBANV_low=5542.48, SD=2753.34, N=56; MBANV_high=7302.13, SD=3943.30, N=55; t=2.73, p=.01). With this information Volkswagen can decide to address less involved consumers primarily

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Brand attitude</th>
<th>Brand familiarity</th>
<th>Purchase intentions</th>
<th>Choice probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>BANV(E,S)<em>j = \sum</em>{a=1}^{m} E_a \times S_a</td>
<td>.22**</td>
<td>.26**</td>
<td>.25**</td>
<td>.16*</td>
</tr>
<tr>
<td>BANV(E,S,F)<em>j = \sum</em>{a=1}^{m} E_a \times S_a \times F_a</td>
<td>.34**</td>
<td>.31**</td>
<td>.33**</td>
<td>.22**</td>
</tr>
<tr>
<td>BANV_j = \sum_{a=1}^{m} E_a \times S_a \times I_a</td>
<td>.40**</td>
<td>.36**</td>
<td>.35**</td>
<td>.24**</td>
</tr>
</tbody>
</table>

** p<.01, * p<.05.

---

Brand association networks

a) Less involved consumers (split-half low involvement, N = 56)

b) Highly involved consumers (split-half high involvement, N = 55)

Notes: The solid circles represent core associations of Volkswagen’s brand image, dashed-line circles represent non-core associations, and the single, double, and triple lines represent the average strength of the association linkage across all respondents. The diameter of each circle represents the average individual importance of an association for the overall product evaluation in a car purchase situation, whereby a larger diameter indicates greater importance. The darkness of each circle represents the average evaluative judgment of an association, with darker circles indicating a more favorable evaluation.

Fig. 3. Brand association networks. a) Less involved consumers (split-half low involvement, N = 56). b) Highly involved consumers (split-half high involvement, N = 55).
through prospective advertising campaigns to enhance the overall brand evaluations of these consumers.

In addition to quantifying the differences in the overall favorability of brand association networks, BANV provides explanations for the differences between two brand association networks. For example, the key first-order brand associations, which are most important for consumers’ purchase decisions (e.g., quality and trustworthiness in Fig. 3a), produce less favorable evaluations in the low involvement segment than the key first-order associations (e.g., quality and satisfaction in Fig. 3b) that emerge for the high involvement segment. Therefore, prospective advertising that targets the low involvement segment should address key brand associations to enhance consumers’ overall network evaluations.

5. Discussion

Existing methodologies measure brand image using brand association networks and have thus uncovered the structure of these networks and provided information about the strength and uniqueness of brand associations. However, existing methodologies, such as the BCM technique, do not reveal explicit measures regarding the favorability of brand associations. This study extends the original BCM approach with explicit information about favorability, that is, an association’s evaluative judgment and individual importance within a purchase situation.

We provide initial evidence that the added information (i.e., advanced BCM) can offer valuable management implications in that it renders the resulting networks more meaningful. We further propose a new metric, brand association network value (BANV), which quantifies the overall favorability of brand association networks using information from the original BCM approach and includes the additional information on the favorability of brand associations, which is collected with Likert scales. Our metric builds on a widely accepted multi-attribute attitude model (Wilkie & Pessemier, 1973) and well-established concepts and theories. With our empirical application, we demonstrate the managerial relevance of BANV and its plausibility in terms of face, nomological, and predictive validity. Overall, our new metric satisfactorily quantifies the overall favorability of consumers’ brand association networks.

Unlike the original BCM approach, the BANV metric can differentiate “good” brand association networks from “bad” ones because it offers a standardized approach for quantifying a network’s overall favorability. The metric also substantially enhances the applicability of the original BCM approach because it can quantitatively compare networks for (groups of) subjects or over time. For example, with the BANV metric, managers can analyze whether specific marketing activities (e.g., brand enrichment, marketing communication, celebrity endorsements, and sponsorships; Keller, 2008) enhance consumers’ brand association networks. Such implications are not available with the original BCM approach, which provides information on the effects on the network structure without combining the information with additional favorability information regarding single brand associations into an overall evaluative judgment of the favorability of the corresponding marketing activity.

Furthermore, BANV not only quantifies the differences in the overall favorability of brand association networks (as simple favorability ratings of the overall strategies might do) but also provides detailed explanations for the differences between the two brand association networks by quantifying the favorability of underlying single brand associations. This step provides valuable implications for managers regarding how to influence consumers’ brand perceptions in terms of specific associations. For example, suppose that a longitudinal study reveals an unfavorable alteration in the overall BANV of brand X, although the overall brand association structure (e.g., first-order associations, core associations, and brand association links) remains unchanged. Information about evaluations of brand associations (i.e., $E_{ij}$) can enable managers to identify which associations deteriorated and caused the unfavorable alteration of the BANV over time. In turn, managers can address those associations promptly with marketing communication activities. Furthermore, information about the importance of brand associations for the overall product evaluation (i.e., $I_{ij}$) reveals which association managers should address through marketing communication if many associations have worsened. The important brand associations can then be addressed primarily through an appropriate allocation of marketing communication resources and ensure a greater impact on consumers’ product evaluations, especially in purchase situations.

Moreover, our new metric is easy to implement and requires limited additional information (i.e., for $E_{ij}$ and $I_{ij}$) in comparison with the original BCM approach. Therefore, the additional time effort demanded of consumers is minimal, and the required number of respondents remains unchanged.

Although our empirical application provides initial evidence that BANV effectively quantifies the favorability of consumers’ brand association networks, several limitations remain that should be addressed by future research. First, we focused our empirical study on merely two brands. Second, the predetermined association set, as well as the associations’ importance information within our advanced BCM procedure, is product category-specific. Thus, the new BANV-metric is primarily applicable to quantitative comparisons of networks regarding one specific brand (e.g., networks of different consumer segments of Volkswagen) and brands of one specific product category (e.g., networks of different car manufacturers).

Furthermore, importance weights for single subjects might be situationally dependent (e.g., single consumers might deem an association to be important for the short-term, whereas they generally would not deem the association to be relevant at all). Therefore, comparisons of BANV values across subjects might suffer from situational dependence bias. One possible way for overcoming this bias is to use an appropriate method of selecting respondents. For instance, if Volkswagen decision makers wonder whether they should target the consumer segment with low involvement or that with high involvement with prospective advertising campaigns (see Section 4.2), the sample should only contain, for instance, consumers from both segments with a substantial interest in purchasing a new car. Furthermore, researchers might test for response styles across subjects, which might be a further source of situational dependence bias (Adler, 1983). Moreover, if brand managers, for instance, are interested in comparing BANV values for two brands, they might use a within-subject design (i.e., each respondent evaluates both brands) to assure that importance weights for both brands are evaluated for the same situation.

Third, we assume that all types of brand associations, including brand-specific attributes and product category associations, are equally weighted. However, product category associations might be shared by the brand’s competitors, whereas brand-specific attributes explicitly differentiate the brand from its competitors in an attempt to generate a sustainable competitive advantage (Fitzsimons, Chartrand, & Fitzsimons, 2008; Keller, 1993; Ries & Trout, 1979). Therefore, further research should consider different (i.e., higher) weights for brand-specific attributes than for product category associations to take these issues into account.

Finally, we based our new metric on the BCM approach because it offers the capability of analyzing brand association networks at an individual disaggregated level. However, because the BANV metric has the aim of quantifying the overall favorability of single brand association networks, it is not limited to the BCM approach and might be extended to other network approaches, such as analytical mapping techniques (e.g., consensus networks within the network analysis). Such analyses provide structural network indices, such as network density, and support comparisons of different networks with respect to consumers’ brand associative structures (e.g., number of existing association linkages), but they do not reveal information about the overall favorability of the underlying association networks (Henderson et al., 1998; 2002). Additional research should determine if the BANV metric might offer a meaningful extension to these approaches as well.
Acknowledgment

The authors thank Michel Clement, Karen Gedenk, Maren Grambeck, Lena Neuner, and Franziska Völckner for their helpful comments on previous versions of this manuscript.

References

Sound symbolism effects across languages: Implications for global brand names

L. J. Shrum a,⁎, T. M. Lowrey a, David Luna b, D. B. Lerman c, Min Liu a

a University of Texas at San Antonio, Dept. of Marketing, One UTSA Circle, San Antonio, TX 78249, USA
b Baruch College, City University of New York, New York, NY 10010, USA
c Fordham University, 113 West 60th Street, New York, NY 10023, USA

1. Introduction

Selecting good brand names for products is a critical step for marketers, and many aspects of a brand name influence brand perceptions. Three experiments investigated the effects of phonetic symbolism (the impact of sound on meaning) on brand name preference, the extent to which these effects generalize to other languages, and the processes that underlie these effects. When choosing brand names, French-, Spanish-, and Chinese-speaking participants who were bilingual in English preferred words in which there was a match between the phonetic symbolism of the words and the product attributes. These results were unaffected by whether participants completed the study in their first or second language, by second-language proficiency, or by whether the Chinese language representations were in logographic or alphabetic form. These findings replicate those of Lowrey and Shrum (2007) and indicate that phonetic symbolism effects for brand name perceptions can generalize across languages, and thus suggest that marketers may be able to embed universal meaning in their brand names.

is based on an entirely different writing system. Consider, for example, the Hydrovive brand in China. The combination of sounds does not map onto the same meanings, or perhaps any meaning, as they do in English and French. In such cases, the marketer must make a choice (Zhang & Schmitt, 2001). One option is to translate the name into Chinese, thus abandoning the sound, to find a name with a similar meaning. The other option is phonetic translation or transliteration, abandoning the meaning to maintain the sound. A third (but more difficult) option is to translate phonosemantically; that is, to translate sound with meaning (Dong & Helms, 2001). Thus, most firms must choose between maintaining the phonetic brand sound and preserving the meaning of the brand name (Francis, Lam, & Walls, 2002; for a review, see Zhang & Schmitt, 2007).

In the examples mentioned, the phonetic qualities pertain to preserving the sound of the name across translations. However, what if the actual sound of the name itself conveys meaning? Moreover, what if the extent of this effect differs across languages? If so, these effects have important implications for considering the sound of the word when constructing new brand names, as well as for the translation strategies that might be adopted. In this study, we investigate this concept and its implications for brand name construction. Numerous studies in psycholinguistics suggest that sounds convey meaning apart from their semantic connotations, a concept referred to as phonetic symbolism or sound symbolism (for a review, see French, 1977). Recent research in marketing has demonstrated that phonetic symbolism has implications for brand name perceptions and preferences (for a review, see Shrum & Lowrey, 2007). However, the extent to which these findings generalize to other languages and writing systems has not been sufficiently addressed, which is clearly

⁎ Corresponding author at: University of Texas at San Antonio, Dept. of Marketing, One UTSA Circle, San Antonio, TX 78249, USA. Tel.: +1 210 458 5374; fax: +1 210 458 6335.
E-mail addresses: lj.shrum@utsa.edu (L.J. Shrum), tina.lowrey@utsa.edu (T.M. Lowrey), david.luna@baruch.cuny.edu (D. Luna), lerma@fordham.edu (D.B. Lerman), min.liu@utsa.edu (M. Liu).

0167-8116/5 – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijresmar.2012.03.002
crucial for applying previous findings to international brand naming contexts.

To address this issue, we report a study that is a replication of previous work (Lowrey & Shrum, 2007), but one whose context is relevant to global brand name construction. Specifically, we investigate the effects of phonetic symbolism across multiple languages, including one with a non-alphabetic writing system (Chinese logographic). Our study uses a bilingual context by testing whether the effects also occur in a consumer’s second language and tests whether the effects vary by second-language proficiency. The international context of the investigation allows us to generalize brand name construction recommendations to global marketing and advertising situations.

2. Theoretical development

2.1. Phonetic symbolism and brand name development

Phonetic symbolism refers to a non-arbitrary relation between sound and meaning. It suggests that the mere sound of a word, apart from its actual definition, conveys meaning. Research supporting the notion of phonetic symbolism has shown that the distinct sounds resulting from different letter combinations are consistently associated with the magnitude of concepts such as size, weight, speed, and hardness, at rates above those predicted by chance (French, 1977). For example, front vowel sounds (such as the [i] vowel sound in pip), in which the tongue is positioned toward the front of the mouth, are associated with perceptions such as smaller, faster, brighter, and harder, whereas back vowel sounds (in which the tongue is toward the back of the mouth, as with the [a] vowel sound in pop) are associated with perceptions such as larger, slower, darker, and softer. Similar associations have also been documented for consonants (Klink, 2000).

Recent research has extended the concept of phonetic symbolism to brand name perceptions and preferences. For example, when presented with fictitious brand names, people perceived names with back vowels to be associated with concepts such as thicker (ketchup), darker (beer), and creamier (ice cream) compared to names with front vowel sounds (cf. Klink, 2000; Yorkston & Menon, 2004). More recent research has expanded these findings to show that brand attitudes and preferences can be enhanced when the fit between the phonetically induced perceptions of a brand name and the product’s attributes is maximized. Lowrey and Shrum (2007) constructed fictitious brand names that varied only by one vowel, which represented the manipulation of the front/back vowel sound distinction (e.g., tiddip vs. toddip). Relative to back vowels, front vowel sounds are perceived to be faster, smaller, sharper, cleaner, and crisper. Consistent with the phonetic symbolism hypothesis, front vowel sound words were preferred over back vowel sound words, by approximately a 2–1 margin, when participants were asked to choose a brand name for a convertible or a knife. However, the opposite was true when participants were asked to choose a brand name for an SUV or a hammer — again by approximately a 2–1 margin.

Although research on phonetic symbolism and brand names suggests that the sounds of brand names influence brand name preferences, there are clear limitations of these studies that inhibit their applicability to international contexts. These limitations include the fact that the majority of research has been conducted only in English and in the United States, has used only alphabetic writing systems, and has not accounted for possible language proficiency effects when the brand name is foreign-sounding or presented in a second-language context.

2.2. Hypotheses

To address these shortcomings in the literature, we conducted a replication of Lowrey and Shrum (2007), but varied a number of factors to test the extent to which the findings generalize across situations that are applicable to international brands. The primary hypothesis we tested is that particular words will be preferred as brand names when the phonetic connotations of the words are consistent with the product attributes. We also varied the language in which the study was presented (English, Spanish, French, and Chinese), whether the language was the first or second language for bilingual speakers, and for Chinese language administrations, whether the writing system was alphabetic or logographic. We also measured language proficiency. For all of these language factors, our expectations were less clear. First, although phonetic symbolism effects have been noted in several languages (Brown, 1958; Sapir, 1929). Second, although fluent and non-fluent speakers process second-language information differently (Luna & Peracchio, 2001; Zhang & Schmitt, 2004), it is not clear whether such processing differences influence phonetic symbolism effects. Third, theorists hold differing views on whether phonetic symbolism effects should be observed for logographic word representations (cf. Chua, 1999; Fang, Horng, & Tzeng, 1986; McCusker, Hillinger, & Bias, 1981; Perfetti & Zhang, 1991).

3. Experiments 1a–1c

3.1. Method

Data collection was conducted in three countries (Experiments 1a–1c) to test our hypotheses. The experiments represented a close replication of Lowrey and Shrum (2007), which crossed vowel sound with product category. Spanish-, French-, and Chinese-speaking participants were fluent in English expressed preferences between brand name pairs that differed only in their primary vowel sound (front vs. back) and did so as a function of product category. In addition, Chinese-speaking participants received brand name stimuli that were constructed using either alphabetic letters or logographic symbols. We also manipulated the languages in which the experiments were completed — whether participants completed the experiment in English or a different language — and we measured their proficiency in the two focal languages.

3.1.1. Participants, procedure, and measures

Participants in Experiments 1a–1c spoke French, Spanish, or Chinese and were bilingual in English. Participants in Experiment 1a (n = 106, 58 women, 47 men, 1 missing; Mage = 23.7 yrs., SD = 2.57) were undergraduates at a French university, participants in Experiment 1b (n = 88, 39 women, 48 men, 1 missing; Mage = 23.6 yrs., SD = 5.53) were undergraduates at a university in the United States with a substantial proportion of Hispanic students, and participants in Experiment 1c (n = 181, 104 women, 77 men; Mage = 31.8 yrs., SD = 7.56) were Chinese participants who were recruited by students in a graduate research course at a university in Taipei.

Participants in all three experiments received the same set of stimuli in the form of questionnaires that differed only in the language in which they were administered. Participants were told that they were participating in a study of brand names. In the first part of the questionnaire, participants were presented with a series of six word pairs (due to translation errors, only four word pairs were used in Chinese logographic conditions). Each word pair differed only by one vowel, which represented the phonetic symbolism manipulation of front versus back vowel sounds. Artificial words were used to avoid semantic associations. Although the artificial words are technically not translatable because they have no meaning, the instructions were translated across languages, a process that was expected to prime that language’s pronunciations and sound associations. The order of presentation was counterbalanced, and all words were evaluated separately by individuals who were bilingual in
English and the target language, to ensure that the pronunciation of the words was as intended and did not closely resemble a real word, which might prime some semantic association. The set of stimuli are shown in Table 1.

Participants were asked to indicate their preferences between each word pair as brand names for a 4×4 vehicle, a hammer, a 2-seater convertible, or a knife. Product categories were pretested to establish that they were properly understood. Because we had similar predictions for the 4×4 vehicle and hammer and for the 2-seater convertible and knife, to conserve statistical power, we combined the product categories so that some participants expressed brand name preferences for both a 4×4 vehicle and a hammer (three word pairs for each; order was randomized) while other participants expressedbrand name preferences for both a 2-seater convertible and a knife (three word pairs for each; order was randomized). This allowed us to collapse across product categories for which back vowel words (4×4 vehicle, hammer) or front vowel words (convertible, knife) were expected to be preferred, if circumstances warranted. Finally, we manipulated the language in which the study was administered: in English, in the language that was the focus of that particular experiment (French, Spanish, or Chinese), and in Experiment 1c, in either Chinese alphabetic or logographic depictions.

Following the brand name preference exercise, participants completed a 13-item language proficiency scale (α=.92) that measured their proficiency in both English and either French, Spanish, or Chinese (Luna, Ringberg, & Peracchio, 2008). Participants also indicated their age, gender, and first language. Finally, they were asked to indicate what they believed the purpose of the study was (none correctly guessed the purpose).

3.2. Results

3.2.1. Effects of sound as a function of product

Our focal hypothesis was that preference for front versus back vowel sound words as brand names will vary as a function of product category: Front vowel sound words will be preferred over back vowel sound words for 2-seater convertible and knife, and back vowel sound words will be preferred over front vowel sound words for 4×4 vehicle and hammer. Thus, we expected a crossover interaction between vowel sound and product. To test these possibilities, we first created continuous dependent variables that represented the proportion of front and back vowel sound words chosen for each product category (e.g., preferring three back vowel words out six=50%).

Preliminary analyses indicated that, as expected, responses did not differ as a function of whether the brand name was for a 2-seater convertible or knife, or as a function of whether the brand name was for a 4×4 vehicle or hammer. Thus, we combined the two pairs to form two product categories: convertible and knife, and 4×4 and hammer. The effects of order and gender were not significant, and thus were not included in the analysis. Next, we combined data from all three experiments into one dataset, and coded experiment as an independent variable. To assess the effects of the alphabetic versus logographic administration in experiment 1c, we coded these factors as two separate experiments for analysis purposes. We then conducted a 2 (vowel sound) × 2 (product category) × 4 (experiment) mixed model analysis of variance (ANOVA), with vowel sound a within-subjects factor and product and experiment between-subjects factors. This analysis allows us to test our overall hypothesis but also determine whether findings differed significantly across experiments.

As predicted, the interaction between vowel sound and product category was significant (F(1, 367)=63.87, p<.001). The preference results as a function of vowel sound and product category can be seen in the top panel of Table 2. Replicating Lowrey and Shrum (2007), front vowel sound words were preferred over back vowel sound words for convertible and knife (58% to 42%; t(188)=5.33, p<.001, one-tailed). In contrast, for 4×4 vehicle and hammer, the predicted opposite pattern was observed: Back vowel sound words were preferred over front vowel sound words (59% to 41%; t(185)=5.37, p<.001, one-tailed). Thus, the predicted crossover interaction was observed.

3.2.2. Effects of language and language proficiency

The three-way interaction between sound, product, and experiment fell just short of significance (F(3, 367)=2.61, p=.052). To decompose this interaction, we performed sound × product ANOVAs for each experiment. The findings from this analysis can be seen in the middle and lower panels of Table 2. The results show that the pattern of effects for the four conditions is consistent: The expected crossover interaction in which majority preference for front versus back vowel sound words changes as a function of product category was observed in each instance (all ps<.006). Individual paired comparisons within product for each experiment indicated that the predicted differences were also significant (all ps<.02, one-tailed), with two exceptions, but both of which were in the expected direction. For Experiment 1c (Chinese alphabetic), the expected preference for front vowel sound words (53%) over back vowel sound words (47%) for 2-seater convertible and knife was not significant (p>.15); for Experiment 1c (Chinese logographic), the expected preference for back vowel

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experiment 1a–1c: brand name preference as a function of vowel sound, product category, and language conditions.</td>
</tr>
</tbody>
</table>

| Product category | % front vowel words preferred | % back vowel words preferred |
| --- |
| All languages combined |  |  |
| Convertible/Knife | 58% | 42% |
| 4×4 SUV/Hammer | 41% | 59% |
| French (Exp. 1a) |  |  |
| Convertible/Knife | 60% | 40% |
| 4×4 SUV/Hammer | 37% | 63% |
| Spanish (Exp. 1b) |  |  |
| Convertible/Knife | 56% | 44% |
| 4×4 SUV/Hammer | 42% | 58% |
| Chinese Alphabetic (Exp. 1c) |  |  |
| Convertible/Knife | 53% | 47% |
| 4×4 SUV/Hammer | 41% | 59% |
| Chinese Logographic (Exp. 1c) |  |  |
| Convertible/Knife | 76% | 24% |
| 4×4 SUV/Hammer | 48% | 54% |

Note. — Comparing across columns, numbers with different superscripts differ at p<.05, one-tailed.
sounds (54%) over front vowel sounds (46%) for 4×4 and hammer was not significant (p > .20).

An inspection of the results also shows that the size of the effects vary somewhat across experiments, which accounts for the three-way interaction. In particular, effect sizes for the French and Chinese logographic conditions tend to be larger than those of the Spanish and English conditions, with the Chinese logographic effect sizes being primarily driven by the front vowel sound effect for convertible/knife category. Further analyses confirmed this observation, with the effect size of the Chinese logographic condition differing from the Spanish (p < .05) and Chinese alphabetic (p < .03) administrations. The difference between the French and the Chinese alphabetic effect sizes approached significance (p < .07).

Finally, we tested whether the effects differed as a function of language proficiency (scale measure) or whether the stimuli were administered in the participants’ first or second language. Neither variable had any significant influence (both Fs < 1 for each interaction), nor did their inclusion alter the interaction between vowel sound and product category.

4. General discussion

Inputs into brand name perceptions are numerous and complex, and a number of factors may influence consumers’ preferences for one brand name over another. In the studies presented here, we showed that the sound of a name, through its phonetic symbolism, is one factor that influences brand name preference. Across three experiments, we showed that preference for a particular brand name over another can be influenced not only by the fit between the name’s phonetic symbolism and the attributes of the product, but in fact the preference as a function of this fit can be reversed. Moreover, we showed that this effect is remarkably stable. We demonstrated the effect in four different languages — English, French, Spanish, and Chinese — and for both alphabetic and logographic language formats in Chinese. We also showed that these effects hold equally for one’s own language and for bilinguals in a second language. For the bilingual conditions, we also showed that this effect does not appear to be affected by language proficiency.

These results add to the growing literature on marketing applications of phonetic symbolism effects. They also provide a theoretical contribution, particularly with respect to the processing of logographic versus alphabetic scripts. The findings suggest that phonetic information is encoded from brands when they are written in logographic scripts, affecting perceptions of those brands, at least when semantic information is encoded from brands when they are written in logographic versus alphabetic scripts. The Focus and the Equinox may violate the front/back vowel sound distinction in the counterexamples we mentioned for naming vehicles. Although some common examples of real brand names consistent with this logic easily come to mind (e.g., Hummer, Tundra (Toyota) for large, powerful vehicles; Prius (Toyota), Twingo (Renault) for small, light vehicles), exceptions are also easily generated (e.g., Ford Focus for a small car, Chevrolet Equinox for a large SUV).

Three important points are worth noting. First, our focus on the front/back vowel sound distinction was primarily to test a theoretical proposition: Does the sound of a brand name influence perceptions and preferences that are generalizable across languages for bilinguals? The decision to use only the front/back vowel sound distinction and hold all other sounds constant was a methodological choice to maximize construct validity. For real brand names, however, the situation is much more complex. The front/back distinction refers to vowel sounds, but there are a number of consonant sounds that have been shown to influence perceptions as well. Examples include fricatives versus stops, and voiceless versus voiced consonants. Moreover, not only are these two sets of categorizations orthogonal (and thus one can have voiced and voiceless fricatives), but some categorizations also have additional dimensions (e.g., occlusive vs. nasal stops). All of these categorizations have been shown to influence perceptions through their sound symbolism (Shrum & Lowrey, 2007).

Thus, the main recommendation that emerges from phonetic symbolism research for brand name creation is that marketers should attempt to maximize the sound-attribute fit. Such fit must be calibrated based on a detailed knowledge of how the sounds of brand names map onto their respective meanings across multiple dimensions. Our point is that knowledge of phonetic symbolism effects would be useful in the brand naming process, both by enhancing sound associations and avoiding bad ones.

The second point we want to stress is that there is much more to a word than just its sound. In fact, sound often plays a very minor role in relation to semantics in constructing brand names. This is evident in the counterexamples we mentioned for naming vehicles. Although the Focus and the Equinox may violate the front/back vowel sound guideline, the names clearly have a meaning, and it is reasonable to assume that semantic connotations will often overwhelm sound connotations. However, when considering two equally attractive brand names that convey meaning through their semantic associations, sound symbolism may provide an added value.

The third point we want to make concerns whether we should expect to see evidence of sound symbolism across brand names for a particular product category, such as the automobile category we chose for our stimuli. The answer depends on a number of variables. One is whether particular product categories tend toward the use of semantics in constructing brand names. In such cases, one might expect to see evidence of the effect only in instances in which the names are fictitious. Although most product categories may rely more heavily on semantics than phonetics, there are also well-known brand names that are made up (Kodak, Exxon).

That said, there are some product categories that may tend to use fictitious names, particularly those likely to use numeric or alphanumeric brand names (Pavia & Costa, 1993). One particular product category that tends, almost entirely, to use fictitious brand names is medication trade names (e.g., Avistan and Taxol, both cancer medications); in fact, there is evidence that phonetic symbolism may be related to the development of brand names in that category.
and Glinert (2008) coded the trade names of 60 frequently used cancer medications in terms of the frequency in which they had voiced or voiceless consonants. They reasoned that because voiceless consonants are associated with concepts such as smaller, lighter, and faster (Klink, 2000; Newman, 1933), medication trade names with voiceless consonants might be associated with more tolerable chemotherapy and would thus be more likely to be used in trade names than voiced consonants. Abel and Glinert (2008) found that this was indeed the case: Voiceless consonants were used in cancer medication brand names more often than would be predicted by their base rate in the English language.

In conclusion, the results of this study support the findings of previous studies that showed that phonetic symbolism influences brand name perceptions and that brand name preference can be enhanced when the fit between the concepts associated with the sound of the brand name and the attributes of the product are maximized. In addition, the results extend previous findings by showing that they generalize to other languages in both alphabetic and logographic writing systems, have similar effects for bilinguals in both their first and second languages, and hold regardless of language proficiency. Thus, an understanding of phonetic symbolism effects represents an additional tool for brand managers when constructing brand names, including names for international brands.

Acknowledgements

We thank the editors and reviewers for their very constructive guidance and encouragement. The paper also benefited from feedback from audiences at the Society for Consumer Psychology and Association for Consumer Research conferences, and from the marketing faculty and doctoral students at HEC Paris. L. J. Shrum and Tina M. Lowrey acknowledge the financial support from UTSA College of Business Summer Research Grants.

References

Negotiating when outnumbered: Agenda strategies for bargaining with buying teams

Charles Patton a,*, P.V. (Sundar) Balakrishnan b,1

a Department of SCM & Marketing Sciences, Rutgers Business School, Rutgers University, 1 Washington Park, Newark, NJ 07102, United States
b School of Business, University of Washington Bothell, MS 358500, 18115 Campus Way NE, Bothell, WA 98011-8246, United States

A R T I C L E   I N F O

Article history:
First received in 31, March 2011 and was under review for 5 months
Available online 23 June 2012

Area Editor: Sönke Albers

Keywords: Negotiations
Agenda
Sales force management
Buying team

A B S T R A C T

The authors empirically investigate how the choice of agenda strategies may enhance economic gain and promote customer relationships when a single salesperson must bargain with a buying team. The authors develop a framework of multi-issue negotiations for examining two key agenda decisions: selecting simultaneous or sequential negotiations; and, within sequential negotiations, determining in which order of importance multiple issues should be bargained. Using face-to-face bargaining settings, the authors demonstrate that, compared to the benchmark of single-buyer vs. single-seller negotiations, simultaneous bargaining of issues with a buying team raises buyers' perceptions of their power and influences a seller's bargaining style. Contrary to conventional wisdom, however, these effects do not disadvantage the single salesperson when tasked with bargaining with a buying team, as the salesperson is no worse off economically than when he or she engages in single-buyer vs. single-seller negotiations. Directly comparing simultaneous to sequential agenda strategies, the authors show that simultaneous negotiations result in more integrative agreements; increased profit to the seller; while at the same time lead to increased satisfaction to the buyers. In sequential negotiations, the ordering of the relative importance of the issues to the parties affects the seller's pre-negotiation disposition, bargaining styles, and—of critical importance to the seller—the likelihood of reaching an agreement. The authors provide managerial implications and contrast them with general beliefs.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Historically, in business markets, negotiations centered on the purchasing agent, an individual tasked with bargaining with salespeople to satisfy the organization's requirements for products or services (Hutt & Speh, 2009). Over the past two decades, however, the buying process among business-to-business customers has been steadily evolving from being primarily the domain of purchasing departments to encompassing the more multi-functional approach of team buying. As Morgan (2001, p. 28) observes, “Cross-functional team buying got its start in the late 1980s when companies began readjusting organizational structures to make them more flexible and competitive.” He found that buying teams are highly popular and in wide use; nearly seventy percent of the companies sampled used or were interested in using team buying and sourcing techniques. Two examples illustrate the broad nature of this transition. Ceparano (1995, p. 24) reported that the purchase of packaging machinery had changed dramatically in the past 10 years with the adoption of buying teams being commonplace. Indeed, at a major packaging machinery exposition, a session was entitled “Team Buying: Do it the Right Way, The profitable Way.” During this same time frame, Liebeck (1996, p. 1) observed that “The traditional 'silo' approach to buying merchandise at Kmart is being dismantled, replaced by a team-buying concept that the giant retailer hopes will improve customer service, in-stocks, merchandise assortments and, ultimately, profitability.”

Under these circumstances, an individual salesperson is solely responsible for negotiating a number of issues, some or all of which fall under the bargaining authority of separate buying team members. Within this context, the salesperson must not only seek successful economic negotiation outcomes but also must balance this objective within the larger context of fostering long-term customer relationships. Given these challenging bargaining environments and complex negotiation goals, we examine approaches salespeople may use in setting their negotiation agendas, a factor long recognized as critical in determining negotiation outcomes (Schelling, 1956).

Agendas are a means of structuring discussions between individuals and groups and comprise the domain of issues along with their ordering for discussion or negotiation. In business markets, negotiation is recognized as the central mechanism to achieve coordination between parties to an exchange (Balakrishnan & Elashberg, 1995; Elashberg, Lilien, & Kim, 1995; Srivastava, Chakravarti, & Rapoport, 2000). These purchases, moreover, account for the majority of the economic activity in industrialized countries (Dwyer & Tanner, 2009). Accordingly, we investigate a number of strategic agenda decisions that are critical for improving a salesperson’s negotiating effectiveness regarding both
short-term gain and long-term relationships with customers (Mantrala et al., 2010; Palmatier, Scheer, Houston, Evans, & Gopalakrishna, 2007).

To better understand which agenda strategies may be most advantageous, we examine the changes that occur in pre-negotiation dispositions and bargaining behaviors when a single seller bargains with a buying team compared to a situation in which a single seller bargains with a single buyer. We find that salespeople should eschew the conventional wisdom that suggests that teams have an advantage (Thompson, 2011). Our research indicates that bargaining with multiple buyers does not necessarily lead to lower profits. Rather, this setting is likely to lead to more integrative agreements, i.e., higher joint profits. Further, we find that bargaining multiple issues simultaneously with all buyers, rather than each issue separately with a single buyer, is likely to increase a salesperson’s profits, buyers’ satisfaction, and the likelihood of reaching an agreement.

We begin by developing a framework to structure the factors salient to agenda setting for negotiation situations in which a single seller must bargain with multiple members of a buying team. Next, we develop two sets of hypotheses related to selecting an agenda under likely buying team negotiation scenarios. We use single-seller vs. single-buyer negotiations as a benchmark to gauge how the team buying scenarios have altered buyers’ and sellers’ perceptions, behaviors, and outcomes. We also conduct a replication of two simultaneous negotiation scenarios and undertake a survey of sales professionals to gain their perspectives. Finally, we provide suggestions for structuring agendas when bargaining with buying teams.

2. Research framework and hypotheses

Our framework comprises four progressive stages and describes the linkages between the key agenda strategies and their negotiated outcomes. Fig. 1 illustrates the framework and depicts the associated hypotheses.

The initial stage, Negotiation Agenda Strategies, depicts two basic strategic agenda decisions regarding multi-issue negotiations. Our research focuses on these two strategic agenda decisions that make up the foundation of a sales agenda. The first strategic decision involves choosing between a simultaneous and a sequential agenda. In a simultaneous agenda, negotiators may bargain all of the issues contemporaneously. In a sequential agenda, negotiators consider the issues singularly and do not reintroduce an issue once they have reached agreement on that issue and have begun to address the next issue (Thompson, Mannix, & Bazerman, 1988). Negotiations under each of these agenda scenarios become more complex when one of the parties is composed of more than a single individual, such as when a single seller bargains not merely with one buyer but with a buying team. In a simultaneous agenda, all members of the buying team and the seller meet together and freely bargain over all issues. In a sequential agenda, a seller meets in succession with each individual buyer to bargain only over those issues that the particular buyer represents.

The second strategic agenda decision arises within sequential negotiations and involves selecting the order in which to discuss multiple issues. While any ordering of the issues is possible in a sequential agenda, two issue orders merit particular attention. As Dobler, Lee, and Burt (1984, p. 223) observe: “most authorities feel that the issues should be discussed in the order of their probable ease of solution” as a means of promoting the overall negotiation process. Therefore, we believe that examining issues in an increasing order of importance may provide insight into factors that promote the negotiation process. Conversely, we believe that examining issues in a decreasing order of importance offers a high probability of uncovering factors that retard the negotiation process.

Fig. 1. Single seller, multiple buyers, multi-issue negotiation process.
The second stage in the framework, Influences on Bargaining Behavior, illustrates key expectations, perceptions and behaviors that influence the agenda strategies. We explore aspirations and power because of their recognized importance and pervasiveness in the body of negotiation literature (e.g., White & Neale, 1994; Wolfe & McGinn, 2005). We examine expectations regarding relationship valence, as these expectations likely influence the manner in which negotiators approach and conduct their bargaining (Weitz, 1981). Lastly, we examine negotiation styles because studies show that they potentially play a critical role in negotiation processes and outcomes (e.g., 2000; Shell, 2001).

The third stage, Negotiation Processes, delineates the simultaneous and sequential ordering of the issues. The fourth and last stage of our framework depicts the various objective and subjective negotiation outcomes. The objective outcomes we investigate are the profits attained by each party; the dyadic-level profits, used to assess the ability to achieve integrative outcomes; and, potentially the most important aspect of negotiations, the likelihood of reaching an agreement. We also examine the affective disposition of the buyers in terms of satisfaction, a key relational outcome of negotiations.

2.1. Hypotheses: sequential vs. simultaneous negotiation agendas

Kim, Pinkley, and Frangale (2005) observe that researchers acknowledge relative power as one of the most important factors in determining the outcomes of negotiated agreements. Research shows that bargainers possessing greater relative power earn higher profits than those in weaker positions. Most studies that employ a one-on-one bargaining context hypothesize power as deriving from either a greater number of alternatives or from knowledge of the other party’s alternatives (e.g., McAlister, Bazerman, & Fader, 1996).

In the context of simultaneous negotiations involving a single seller and a team of buyers, we posit that power derives directly from the composition of the parties. Both the buyers and the seller are likely to perceive that there is strength in numbers on the part of the buyers, i.e., two heads are better than one (Perkins, 1993). Thus, the presence of several buyers at the bargaining table during simultaneous negotiations is likely to generate raised perceptions of their relative power. In this regard, our conceptualization of power in negotiations is consistent with Wolfe and McGinn (2005) who view power as a perceived and relational construct.

H1a. Buyers who are part of a buying team engaged in simultaneous negotiations with a single seller will have higher perceptions of their power than buyers engaged in single-buyer vs. single-seller negotiations. Conversely, buyers who bargain individually as part of a buying team engaged in sequential issue negotiations with a single seller will have similar perceptions of power as buyers engaged in single-buyer vs. single-seller negotiations.

Empirical research (e.g., Tajfel, 1970) shows that people relating their experiences with groups and with individuals state that their relationships with groups were more competitive (competitive situations are characterized by negative goal interdependence—one person wins, the others lose). Similarly, when people assess how competitive their relationships would be with groups vs. individuals, they assume that groups will be more competitive (Insko & Schopler, 1998; Pemberton, Insko, & Schopler, 1996). Moreover, when individuals encounter groups, this assumption engenders distrust within individuals, which then drives competitive behavior (Insko et al., 1987; McCallum et al., 1985). Thus, faced with bargaining with a group, a seller is likely to counter the perceived greater competitiveness of the buyers by engaging in a competitive style of bargaining.

H1b. A seller engaged in simultaneous negotiations with a buying team will employ more of a competitive negotiation style than a seller engaged in single-buyer vs. single-seller negotiations. Conversely, a seller who bargains issues sequentially with individual members of a buying team will not employ more of a competitive negotiation style than a seller engaged in single-buyer vs. single-seller negotiations.

Compared to sequential negotiations, simultaneous negotiations naturally allow more opportunity for integrative agreements. Under simultaneous negotiations, however, the raised perceptions of buyers’ power and the more competitive nature of the seller’s bargaining behavior are factors that likely retard effective bargaining and limit integrative outcomes. Studies show that raised levels of power by one of the parties inhibit conflict resolution (Lawler & Yoon, 1993) and lead to less integrative agreements (Wolfe & McGinn, 2005). As Mannix, Thompson, and Bazerman (1989, p. 510) note, under unequal power situations, negotiators “focus on the norms of distribution rather than on ways in which the joint outcomes might be increased.”

The composition of the buying team makes these factors less likely to be dominant under simultaneous negotiations. Studies investigating small group vs. individual problem solving indicate that, on intellective tasks (where there are demonstrably correct solutions), groups tend not only to outperform the average individual, but perform at a level similar to the best performance of an equivalent number of individuals (Bonner, Baumann, & Dalal, 2002). Further, Laughlin, Bonner, and Miner (2002) found that groups outperformed even the best comparison individuals. This superiority of group performance over individual performance is attributed to groups’ superior abilities in information processing (Hinsz, Tindale, & Vollrath, 1997). Thus, under simultaneous negotiations, the trial and error process of bargaining should allow buying teams to generate a heightened perspective of the opportunity for integrative agreements.

H1c. In simultaneous negotiations between a buying team and a single seller, economic outcomes will be more integrative (i.e., result in higher joint profits) than in sequential negotiations between a buying team and a single seller.

Aspiration levels—defined by Pruitt (1981) as a negotiator’s drive for achievement and the levels of utility for which the negotiator is striving—are one of the major constructs employed in the negotiation literature. Empirical research (e.g., White & Neale, 1994) demonstrates that higher aspiration levels result in larger profits for the associated bargainers. Under either of the negotiation scenarios, the single seller must negotiate all the issues. Therefore, he or she likely sets aspiration levels in accordance with his or her perception of the issues’ relative importance. How strongly the seller strives to succeed regarding these issues during bargaining thus will be roughly proportional to each issue’s perceived relative importance. For buyers under simultaneous negotiation, this relationship between aspiration level and issue importance may also exist.

For buyers under sequential negotiations, however, aspirations are unlikely to be proportional to the importance of the issues. Because only one of the buyers bargains each issue during a separate session, each issue is this individual’s sole responsibility during negotiations and thus takes on an added salience (O’Connor, 1997). As Thompson et al. (1988, p. 88) observe, “Explicit issue-by-issue agendas shift the focus of negotiation from the perception of group gain to the perception of winners and losers on each issue.” That is, the issue takes on a level of ego involvement that Balakrishnan, Patton, and Lewis (1993, p. 647) define as “a bargainer’s perception of a close association between certain issues and his or her self-esteem.” Accordingly, a buyer tasked with bargaining an issue individually in sequential negotiations will possess higher aspirations and be less inclined to accept lower profits than if he or she were to negotiate that issue in concert with the other buying team members in simultaneous negotiations.

H1d. Buyers who bargain individually as part of a buying team engaged in sequential issue negotiations with a single seller will
have higher aspirations than buyers engaged in single-buyer vs.
single-seller negotiations. Conversely, buyers who are part of a buy-
ing team engaged in simultaneous negotiations with a single seller
will have similar aspirations as buyers engaged in single-buyer vs.
single-seller negotiations.

Because higher aspiration levels typically lead to higher utilities
(e.g., Zetik & Stuhlmacher, 2002), sellers are likely to be at a disad-
vantage when bargaining under sequential negotiations compared to
simultaneous negotiations.

H1e. A seller engaged in simultaneous negotiations with a buying
team will obtain higher profits than a seller engaged in sequential ne-
gotiations with a buying team.

Buyers' satisfaction with negotiations is critical; researchers find
that the levels of satisfaction with an agreement may affect the de-
sire for continued contact and cooperativeness between the parties
(e.g., Heide & Miner, 1992; Thompson, 1993). Under the simulta-
neous negotiation process, buyers negotiate in a mutually supportive
environment that places the buyers in a position of feeling relatively
more powerful than the seller (H1a). Coupled with the basic percep-
tion that “two heads are better than one” (e.g., Thompson, 2011), buyers are likely to believe that they have thought out a more thor-
ough course of action and achieved a better payoff than the single
seller is capable of achieving. Because simultaneous negotiations
also promote more integrative agreements (e.g., Pruitt, 1981), buyers' satisfaction should be greater under simultaneous negotia-
tions than under sequential negotiations.

H1f. Buyers who are part of a buying team engaged in simultaneous
negotiations with a single seller will have higher satisfaction than
buyers who are part of a buying team engaged sequential issue nego-
tiations with a single seller.

2.2. Hypotheses: issue order within sequential negotiation agendas

Prior to negotiations, buyers and seller are likely to believe that
bargaining from least to most important represents the most benefi-
cial ordering of the issues (Fershtman, 1990). Due to the increasing
importance of the issues, the parties expect to compensate on the
next issue for any shortfalls in the current bargaining. Additionally,
bargainers may learn from the bargaining experience, and thus in-
creasing skill can be used to advantage on the more important issues
(e.g., Thompson, 1990). In contrast, when issues are bargain in the
order of most-to-least important, there is no benefit from being able
to learn while bargaining on the issues of lesser importance. Further,
compensating on later issues for achieving less profit than desired on
previous issues becomes less likely.

Thus, a single seller bargaining issues in the order of most-to-least
important is likely to expect greater difficulty in achieving his or her
desired outcomes. Under this bargaining condition, a seller is also likely
to feel that he or she cannot afford to be as cooperative and that the
opposing buyer will also be less cooperative. In this regard, studies
find that people who expect others to cooperate are themselves more
likely to cooperate, and vice versa (e.g., Messick & Brewer, 1983;
Wiener & Doescher, 1994). In turn, research shows that where the
parties stand on a “cooperative and friendly” continuum are important
determinants of negotiation processes and outcomes (e.g., Halpern,
this aspect of a relationship the “valence” of the relationship.

H2a. A seller engaged in sequential negotiations with a buying team
will have less positive expectations of relationship valence when
bargaining issues in the order of most-to-least important than a seller
bargaining issues in the order of least-to-most important.

Following directly from the arguments for hypothesis 2a, the seller’s
perception of less positive relationship valence and the circumstance
of bargaining in the order of most-to-least important are likely to shape
the seller’s choice of negotiation style (Ganesan, 1993). That is, forcing
a seller to bargain the issues in the order of most-to-least important
has placed the seller at a perceived disadvantage. This disadvantage
should instill in the seller a desire to mitigate the less positive environ-
ment to move through the negotiations and bring about an agreement.
This desire should engender the use of less aggressive negotiation styles
to maximize the opportunity to conclude an agreement.

H2b. A seller engaged in sequential negotiations with a buying team
will make greater use of Avoidance and Yielding negotiation styles
when bargaining issues in the order of most-to-least important than a
seller bargaining issues in the order of least-to-most important.

The two sequential negotiation scenarios also have decidedly
different likelihoods of reaching agreement. Buyers are driven by
raised aspiration levels in sequential negotiations (H1b). Buyers also
may want to make a deal to avoid disappointing others in their party or to preclude the remaining members of the buying team
from bargaining. However, sellers answer only to themselves. When a
single seller bargains under the scenario in which issues are
bargained in order from most-to-least important, the seller enters the
negotiations expecting a more difficult environment (H2a), which neg-
atively impacts his/her bargaining stance. During bargaining, the raised
aspiration levels of buyers under sequential negotiations are likely to
reinforce this expectation. If the seller perceives that bargaining on
the initial, important issues will not yield satisfactory outcomes, the
seller may have little desire to continue the negotiation process under
these conditions, as the remaining issues offer diminishing opportuni-
ties to recoup the perceived shortfall.

H2c. A seller engaged in sequential negotiations with a buying team
has a greater likelihood of reaching an agreement when bargaining
issues in the order of least-to-most important than a seller bargaining
issues in the order of most-to-least important.

3. Empirical investigations

3.1. Methodology

Our goal was to design experimental negotiation scenarios that
would provide an accurate representation of a seller bargaining with a
buying team under simultaneous and sequential agenda strategies.
We considered several key design issues: the number of bargainers
that should comprise the buying teams; the types and number of issues
over which the buyers and sellers should bargain and their associated
bargaining roles; and, the nature of the negotiation processes that the
experimental negotiation scenarios should represent.

For each of the multiple buyer scenarios, we used three individuals
to represent the composition of the buying teams. We chose this num-
ber for several reasons. First, and most importantly, three buyers reflect
typical buying-center size (McWilliams, Naumann, & Scott, 1992). Sec-
ond, three individuals enabled us to extend the investigations beyond
two-person groups, which have been used in previous limited team ne-
gotiation research (e.g., Brodt & Tuchinsky, 2000). Third, three individu-
als are more commonly used in a number of group research contexts
(Morgan & Tindale, 2002). Finally, recent research by Laughlin, Hatch,
Silver, and Boh (2006, p. 648) suggests “that 3-person groups are neces-
sary and sufficient to perform better than the best individuals on highly
intellective problems.”

Selecting three buyers as the appropriate numerical representation
of a buying team also allowed us to determine that three issues should
form the basis of the negotiations. That is, each buyer is responsible for a
separate issue. We selected issues that we believed were realistic and
would be easily understood by the participants. Namely, we selected three issues to negotiate for a new clothing line: retail margins, advertising support and credit terms. Participants played the role of either the Marketing Manager of a clothing manufacturer or a buying team member (Advertising Manager, Chief Buyer, or Director of Finance) for a large retailing organization.

In the commercial scenario we selected, we chose to express issue importance in terms of greater financial consequence (i.e., monetary profits). We thus avoided non-monetary issues (such as being “environmentally friendly”) on which parties may not agree because of differing views regarding these issues’ relative “importance.” Two of the issues (retail margin and credit terms) had diametrically opposing importance for the parties. Specifically, retail margin had high profit potential (most importance) to the seller but low profit potential (least importance) to the buyers. On the other hand, credit terms had high profit potential (most importance) to the buyers but provided low profit potential (least importance) to the seller. Thus, these two issues offered the opportunity for trade-offs and the development of integrative bargaining solutions. Consequently, in the experimental design, the critical aspect for sequential negotiations involved the order in which the parties negotiated the issues of credit terms and retail margin. The issue of advertising support was purely distributive in nature, as it was of equal importance to both parties, and we always employed it as the middle issue. Moreover, the use of this distributive issue as a “filler” made the bargaining task a little more complex as it prevented easy discovery of mutually beneficial solutions.

We selected four distinct negotiation scenarios as representative of our agenda strategies (Fig. 2). Treatment A employed a sequential negotiation scenario in which the single seller negotiated individually with each of three buyers over one issue at a time. The negotiations did not progress to the next issue until the parties reached agreement on the current issue. The parties bargained the issues according to the seller’s least-to-most important. Treatment B used similar methods to Treatment A, but the parties bargained the issues in the reverse order of importance: most-to-least important. Treatment C employed a simultaneous agenda scenario in which the single seller negotiated with all three buyers at the same time and the parties raised and bargained issues at their discretion. Treatment D was similar to Treatment C with the major exception that only a single buyer was responsible for negotiating all three issues. Treatment D played the important role of acting as a benchmark by which to gauge the changes in buyers’ and sellers’ pre-negotiation expectations and bargaining behaviors when engaged in buying team negotiations.

In the sequential negotiation scenarios (Treatments A and B), we chose the ordering of issue importance to explore the oft-stated advantageous strategy of bargaining issues in the order of least-to-most important. Conversely, we explored the opposite extreme via bargaining issues in the order of most-to-least important to examine which ordering of issues should logically follow as the least advantageous issue order. Both these issue orderings maintained the integrative potential of the bargaining task. Finally, the use of opposite issue orderings enabled an experimental design that allowed us to compare the combined sequential scenarios to the multiple-buyer vs. single-seller simultaneous negotiations and the one-on-one negotiations.

The sample consisted of upper class university students majoring in business administration. We conducted the negotiation experiments over a period of several months. As small groups of participants became available through the recruitment process, we assigned a specific date and time to report to a designated meeting room. Upon participants’ arrival, we randomly grouped them into the required negotiation roles using one of the four negotiation scenarios. We seated sellers and buyers in separate rooms and then gave each participant a specific packet of materials. These materials provided important information: 1) an explanation of the task, the participant’s role, and the nature of the negotiations; 2) instructions regarding how and when to fill out the pre- and post- negotiation questionnaires, which collected necessary information about their perceptions, bargaining styles, and outcomes; and, 3) a payoff table listing the profits ($ in millions) that would accrue from bargaining over each of the three issues. The payoff table and the task information were private information specific to each role. The negotiation scenarios (see Web Appendix) and payoff tables (Appendix A) we used represented a variation of those used by Patton and Balakrishnan (2010). Note that the most integrative agreement available was $104 million. In contrast, a distributive outcome generated only $80 million of total profits. Thus, it was possible to expand the size of the bargaining pie by as much as thirty percent.

We gave the buyers and sellers time to familiarize themselves with the materials and indicate that they understood the task. We instructed buyers and sellers that all forms of communication between them were permissible as long as they did not physically share their payoff table with the other party and as long as they considered only the options listed. We allowed buying team members time to caucus and set an overall strategy prior to the first stage of negotiations. At this point, buyers and sellers completed their pre-negotiation questionnaire. Under each experimental treatment, we brought relevant bargainers together in a room to commence negotiations. The study imposed no explicit time limits on the negotiation sessions. However, in both of the sequential negotiation scenarios, we allowed only one member of the buying team to be present with the seller during the negotiation. Once negotiations commenced, we did not permit discussion among the buyers. In the simultaneous scenario, all three buyers were present in the negotiation room with the single seller. As each buyer or seller completed his or her part of the negotiation session, we directed that participant to a separate room to complete the post-negotiation questionnaire.

Two successive pretests of the negotiation instruments minimized the possibility of ambiguous wording. Using the revised instrument, we obtained data from 192 bargainers, resulting in 11 agreements in Treatment A, 11 agreements and 5 non-agreements in Treatment B, 12 agreements and 1 non-agreement in Treatment C, and 15 agreements and 1 non-agreement in Treatment D (Tables 1 and 2).

3.2. Measures

Prior to the commencement of negotiations, we measured buyers’ and sellers’ aspiration levels and expectations concerning their bargaining relationships. We used a pre-negotiation form to assess aspirations by asking the participants to indicate their very best, most likely, and worst acceptable aspirations of profit. We collected each participant’s confidence in these judgments, i.e., the likelihood of attaining each aspiration on a scale of 0 to 100 (Balakrishnan et al., 1993; White & Neale, 1994). We computed aspiration levels as the weighted average of these three expectation judgments and their associated confidence judgments. We assessed expectations concerning relationship valence (Iacobucci & Ostrom, 1996) based on the work of Wish, Deutsch, and Kaplan (1976), which identified four basic dimensions of interpersonal relations. We generated questions mirroring the three items Wish et al. (1976) discovered as having the highest principal component weights for the “positive-negative interpersonal disposition” dimension. That is, we identified the parties’ positions on a “cooperative and friendly” continuum. Each question used a seven point Likert-type scale with the respective anchors ranging from Difficult (1) to Cordial (7), Uncooperative (1) to Cooperative (7), and Antagonistic (1) to Friendly (7). Averaged

---

2 We thank a reviewer for making this observation.

http://faculty.washington.edu/sundar/TechAppendix/Appendix-Agenda-Setting-IJRM.pdf
Buyers had significantly raised perceptions of their relative power compared to buyers engaged in one-on-one (1v1) bargaining (Multi-Buyers Sim. 3v1: 62.8 vs. Single-Buyer 1v1: 51.5, p = .001). In contrast, under sequential bargaining (Seq), buyers’ perceptions of power did not rise compared to in one-on-one bargaining (Multi-Buyers Seq. 3v1: 53.8 vs. Single-Buyer 1v1: 51.5, p = .286). Faced with the perceived greater power of the multiple buyers in simultaneous negotiations, single sellers made far greater use of a Competitive negotiation style than sellers did under one-on-one bargaining (Seller Sim. 1v3: 66.7% vs. Seller 1v1: 13.3%, p = .008). Under sequential negotiations, sellers demonstrated no change in the use of a Competitive negotiation style compared to sellers in one-on-one bargaining (Seller Seq. 1v3: 36.4% vs. Seller 1v1: 13.3%, p = .121). Thus, the findings supported hypotheses 1a and 1b.

In terms of economic outcomes, the average joint profits for multiple buyers and single sellers engaged in simultaneous negotiations were significantly higher by $8.5 million than when they were engaged in sequential negotiations (Sim: $88.0 million vs. Seq: $79.5 million, p = .000). Moreover, the simultaneous average joint profits were also larger than the amount generated in one-on-one bargaining (Sim: $88.0 million vs. 1on1: $82.0 million, p = .036). The simultaneous agreements also may be considered integrative in nature, as their average joint profits of $88 million were significantly different from the $80 million of simple distributive agreements (p = .001). Thus, the findings supported hypothesis 1c (Table 3).

Investigating hypothesis 1e, we found that single sellers’ profits under sequential negotiations were significantly different from single sellers’ profits under simultaneous negotiations (Seller Seq. 1v3: $37.7 million vs. Seller Sim. 1v3: $43.9 million, p = .045). The study

### 3.3. Analysis and results

#### 3.3.1. Analyses: sequential vs. simultaneous negotiation agendas

When engaged in multi-buyer simultaneous negotiations (Sim), buyers significantly raised perceptions of their relative power compared to in one-on-one (1v1) bargaining (Multi-Buyers Sim. 3v1: 62.8 vs. Single-Buyer 1v1: 51.5, p = .001). In contrast, under sequential bargaining (Seq), buyers’ perceptions of power did not rise compared to in one-on-one bargaining (Multi-Buyers Seq. 3v1: 53.8 vs. Single-Buyer 1v1: 51.5, p = .286). Faced with the perceived greater power of the multiple buyers in simultaneous negotiations, single sellers made far greater use of a Competitive negotiation style than sellers did under one-on-one bargaining (Seller Sim. 1v3: 66.7% vs. Seller 1v1: 13.3%, p = .008). Under sequential negotiations, sellers demonstrated no change in the use of a Competitive negotiation style compared to sellers in one-on-one bargaining (Seller Seq. 1v3: 36.4% vs. Seller 1v1: 13.3%, p = .121). Thus, the findings supported hypotheses 1a and 1b.

In terms of economic outcomes, the average joint profits for multiple buyers and single sellers engaged in simultaneous negotiations were significantly higher by $8.5 million than when they were engaged in sequential negotiations (Sim: $88.0 million vs. Seq: $79.5 million, p = .000). Moreover, the simultaneous average joint profits were also larger than the amount generated in one-on-one bargaining (Sim: $88.0 million vs. 1on1: $82.0 million, p = .036). The simultaneous agreements also may be considered integrative in nature, as their average joint profits of $88 million were significantly different from the $80 million of simple distributive agreements (p = .001). Thus, the findings supported hypothesis 1c (Table 3).

Investigating hypothesis 1e, we found that single sellers’ profits under sequential negotiations were significantly different from single sellers’ profits under simultaneous negotiations (Seller Seq. 1v3: $37.7 million vs. Seller Sim. 1v3: $43.9 million, p = .045). The study
largely supported the rationale for lower single sellers’ profits under sequential compared to simultaneous negotiations (hypothesis 1D), namely, the raised aspiration levels of the multiple buyers. The pre-negotiation measures of multiple buyers’ aspiration levels were significantly greater than buyers’ aspiration levels under one-on-one negotiations. However, for the most important issue under simultaneous negotiations, buyers’ aspiration levels were also raised compared to buyers’ aspiration levels under one-on-one bargaining (Multiple Buyers Seq. 3v1: $23.5 million vs. Buyer 1v1: $20.2 million, p = .037).

Examining multiple buyers’ satisfaction (hypothesis 1F), we found greater buyer satisfaction under simultaneous negotiations than under sequential negotiations (Multiple Buyers Seq. 3v1: 5.00 vs. Multiple Buyers Seq. 1v1: 4.41, p = .025). Thus, hypothesis 1F was supported.

3.3.2. Analyses: issue order within sequential negotiation agendas

Examining the pre-negotiation dispositions of the parties (hypothesis 2A), the study found (Table 4) that in Treatment B (seller: most-to-least), single sellers expected the relationship valence (Coefficient Alpha = .751) to be less positive compared to single sellers’ expectations of relationship valence under one-on-one negotiations (Seller Seq. 1v3 Treat. B: 4.00 vs. Seller 1on1: 5.09, p = .003). In contrast, the findings demonstrated no differences in the pre-negotiation expectations between single sellers in Treatment A (Seller: Least-to-Most) and single sellers in one-on-one bargaining (Seller Seq. 1v3 Treat. A: 4.51 vs. Seller 1on1: 5.09, p = .06).

Post-negotiation assessments of bargaining styles (hypothesis 2B) revealed that single sellers in Treatment B (Seller: Most-to-Least), compared with sellers bargaining one-on-one, believed that they far more frequently used both an Avoidance style (Seller Seq. 1v3 Treat. B: 90.9% vs. Seller 1on1: 20.0%, p = .000) and a Yielding style (Seller Seq. 1v3 Treat. B: 100% vs. Seller 1on1: 46.7%, p = .000) in their bargaining. Further, we found no change in the use of the two bargaining styles when single sellers in Treatment A (Seller: Least-to-Most) were compared to sellers engaged in one-on-one bargaining: Avoidance style (Seller Seq. 1v3 Treat. A: 27.3% vs. Seller 1on1: 20.0%, p = .664) and Yielding style (Seller Seq. 1v3 Treat. A: 18.2% vs. Seller 1on1: 46.7%, p = .131).

Investigating hypothesis 2C, we found that 5 of the 16 dyads in Treatment B (Seller: Most-to-Least) failed to reach agreement, while every group negotiating under Treatment A (Seller: Least-to-Most) did arrive at an agreement (p = .000). Furthermore, on examining all of the non-agreement responses, we found in each case that bargainers failed to reach agreement on the first of the three issues, i.e., the most important issue to the single seller and the least important to the buyer. Thus, both the likelihood of reaching an agreement under the two sequential negotiation strategies, hypothesis 2C, and the rationales for these likelihoods, hypotheses 2A and 2B, were supported.

We also ran Mann-Whitney analyses to examine the findings using a nonparametric technique. We confirmed all hypotheses except for hypothesis 2A. We found that a single seller’s expectations of relationship valence, bargaining from least-to-most-important issue, now statistically differed from a seller bargaining under one-on-one negotiations (Seller Seq. 1v3 Treat. A: 4.51 vs. Seller 1on1: 5.09, p = .039).
Because we collected data over several months, we tested for the possibility of cross-talk between study participants. We divided the data from each of the five study treatments (A, B, A&B, C, and D) chronologically into first- and second-half participants. We compared buyers’ and sellers’ profits and satisfaction levels between the first and second half participants. We found that no second half output was larger than a first half output by a statistically significant amount. Therefore, it is unlikely that initial participants provided later participants information that allowed them to better understand the integrative nature of the bargaining, generate greater profits, or attain higher levels of satisfaction.

3.4. Validation

To provide additional support for the findings, we undertook a replication (Kayande, De Bruyn, Lilien, Rangaswamy, & van Bruggen, 2009) of the two simultaneous negotiation scenarios, namely Treatment C (single seller vs. multiple buyers) and Treatment D (single seller vs. single buyer), using MBA students. The validation sample consisted of 82 students pursuing their MBA degrees while working full-time. Their average age was 31.4 years; 66% were male; and they possessed on average 9.5 years of work experience.

We found that buyers engaged in the multiple-buyer simultaneous negotiation scenario had significantly raised perceptions of their relative power compared to buyers engaged in one-on-one bargaining. (Multiple Buyers Sim. 3v1: 60.5 vs. Buyer 1v1: 50.7, p = .002). Single sellers faced with bargaining with multiple buyers met this challenge by making greater use of a Competitive negotiation style than sellers did under one-on-one bargaining (Seller Sim. 1v3: 61.5% vs. Seller 1on1: 20.0%, p = .015). Thus, we find further support for hypotheses 1a and 1b. We also found support for hypothesis 1d, as none of the pre-negotiation measures of multiple buyers’ aspiration levels significantly differed from buyers’ aspiration levels under one-on-one negotiations.

Additionally, we undertook a survey of sales professionals to gain their perspectives on a number of the fundamental expectations and beliefs that form the basis of our research. The sample consisted of 52 field salespeople who were employed by divisions of Fortune 1000 corporations and engaged in business-to-business selling. The average age of the respondents was 41.3 years; 84% were male; and they possessed on average 9.5 years of work experience.

Two findings in particular further confirm the dramatic changes salespeople expect when facing a buying team rather than a single buyer. First, by a three to one ratio, salespeople would prefer to bargain with a single buyer than with a three-buyer team (Single buyer: 75% vs. Multiple Buyers; 25%, p = .000 vs. H0: .50). Second, regarding relative power, salespeople believe there is little difference between themselves and a buyer when bargaining with a single buyer (52.6 points to seller, vs. 47.4 points to buyer, p = .233). In contrast, salespeople believe they will be less powerful than buyers when bargaining with a three-buyer team (39.8 points seller vs. 60.2 points to buyers, p = .000). Details of the validation findings and survey instruments can be found in the web appendix.

4. Discussion

Researchers recognize that little is known about how teams negotiate (O’Connor, 1997). Prior team negotiation research, moreover, only examines bargaining from a simultaneous basis in which bargainers prepare a strategy and then engage in a single bargaining session (e.g., Brodt & Tuchinsky, 2000; Thompson et al., 1988; Thompson, Peterson, & Brodt, 1996). In contrast, we examine buying team negotiations from a simultaneous basis, from a sequential basis, from the reference point of single buyer vs. single seller one-on-one bargaining, and from the more realistic and complex perspective of three person teams.

We find that both the seller’s and buyers’ expectations, dispositions, and behaviors dramatically change when a seller bargains with a buying team rather than in the more common single seller vs. single buyer negotiations. However, we show that these changes depend on whether the buying team members bargain together simultaneously as a group or bargain individually in a sequential fashion. If members bargain simultaneously, buyers raise their perceptions of their own power, while a seller bargains in a more competitive manner. If, however, the
buying team members bargain sequentially, neither condition eventuates. Rather, buyers raise their aspiration levels.

Our findings show that buyers' perceived power and a seller's competitive negotiation style do not inhibit a seller from making greater profits and achieving more integrative agreements when bargaining with a buying team under simultaneous compared to sequential negotiations. Indeed, these factors appear to be trumped by the inherent capacity for greater information processing and exchange by the multiple buyers. Moreover, simultaneous negotiations with a buying team are also likely to be more integrative than single-seller vs. single-buyer one-on-one negotiations. Choosing between a simultaneous and a sequential agenda strategy creates major differences in outcomes, though by quite different mechanisms.

Integrative agreements in simultaneous negotiations appear to occur because the superior information processing ability of multiple buyers allows them to better recognize and take advantage of the tradeoffs they can make (Rangaswamy & Shell, 1997). Accordingly, integrative agreements are more likely to directionally raise a seller's profit than one-on-one negotiations are. In contrast, buyers' aspiration levels which are raised in sequential negotiations relative to one-on-one negotiations are more likely to directionally lower a seller's profit in the sequential negotiations. Together, these opposing directional changes create a significant difference in a seller's profits between the two agenda strategies. For buyers, these same mechanisms are likely to raise profits directionally higher under both negotiation scenarios when compared to one-on-one bargaining. Consequently, these arguments suggest that there will be no difference in buyers' profits between sequential and simultaneous negotiations. Lastly, buyers' satisfaction is higher under simultaneous negotiations than under sequential negotiations because buyers are likely to believe that their greater numbers give them the capacity for superior results and because their profits are directionally higher under integrative agreements.

We must also note a finding that is counter to what we expected. Buyers have raised aspiration levels on the most important issue in simultaneous negotiations (H1d). However, a straightforward rationale may exist for this finding. The buyer charged with the most important issue knows that success of the negotiations hinges largely on the profit obtained on this issue. Therefore, the buyer may feel added pressure to succeed and thus raises his or her level of aspirations.

In contrast to simultaneous negotiations, a sequential negotiation of the issues is likely to generate profits that are the same as in simple distributive agreements. Further, the order of importance in which the issues are bargained may severely impact negotiation impasse rates. If a seller bargains the issues in the order from least-to-most important (Treatment A), the likelihood of reaching an agreement is much greater than when bargaining the issues in the reverse order (Treatment B).

Failure to reach agreement appears to arise because a seller, bargaining under the most-to-least important ordering of issues, expects the relationship valence to be less positive than under one-on-one negotiations and is faced with the most difficult negotiation task first. These conditions also appear to manifest in a single seller's greater use of Avoidance and Yielding negotiation styles compared to in one-on-one negotiations. Because the first negotiation issue is likely to be viewed by a seller as a make-or-break issue in terms of the profit to be received, the seller may have a tendency to discontinue negotiations when difficulties arise. Thus, our examination provides insight into the critical problem of why negotiators fail to reach agreement (Bazerman & Carroll, 1987).

Finally, the results appear to indicate that the order of importance of issues to the buying team in sequential negotiations has little impact on the bargaining process. Under either sequential scenario, buyers have the same levels of perceived power, expectations of relationship valence, bargaining styles, and raised aspirations. Buyers also

---

### Table 3

Analysis of simultaneous vs. sequential negotiation agendas.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Mean</th>
<th>S. D.</th>
<th>Prob.</th>
<th>Hypothesis supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a: Buyers' perceptions of power — (0–100 pts)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-on-one neg. (D) vs. Sequential neg. (A and B)</td>
<td>62.83</td>
<td>12.55</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>Simultaneous neg. (C)</td>
<td>62.83</td>
<td>12.55</td>
<td>.001</td>
<td>Yes</td>
</tr>
<tr>
<td>1b: Single seller's competitive negotiation styles — (yes – no, percent yes)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-on-one neg. (D) vs. Sequential neg. (A and B)</td>
<td>86.36%</td>
<td>10.26%</td>
<td>.121</td>
<td></td>
</tr>
<tr>
<td>Simultaneous neg. (C)</td>
<td>63.60%</td>
<td>13.89%</td>
<td>.008</td>
<td></td>
</tr>
<tr>
<td>1c: Joint profits — ($ MM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential neg. (A and B)</td>
<td>79.45</td>
<td>3.54</td>
<td>.000</td>
<td>Yes</td>
</tr>
<tr>
<td>Simultaneous neg. (C)</td>
<td>88.00</td>
<td>9.17</td>
<td>.000</td>
<td>5 of 6 parts</td>
</tr>
<tr>
<td>1d: Multiple buyers' aspiration levels — ($ MM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Least important issue</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-on-one neg. (D) vs. Sequential (A and B)</td>
<td>5.00</td>
<td>0.71</td>
<td>.025</td>
<td></td>
</tr>
<tr>
<td>Simultaneous (C)</td>
<td>5.00</td>
<td>0.71</td>
<td>.025</td>
<td></td>
</tr>
<tr>
<td>Most important issue</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One-on-one neg. (D) vs. Sequential (A and B)</td>
<td>8.43</td>
<td>1.05</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Simultaneous (C)</td>
<td>9.04</td>
<td>2.64</td>
<td>.422</td>
<td></td>
</tr>
<tr>
<td>1e: Single seller's profits — ($ MM)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential neg. (A and B)</td>
<td>79.45</td>
<td>3.54</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Simultaneous neg. (C)</td>
<td>88.00</td>
<td>9.17</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>1f: Multiple buyers' satisfaction — (1 extremely dissatisfied – 7 extremely satisfied)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential neg. (A and B)</td>
<td>4.41</td>
<td>.85</td>
<td>.761</td>
<td></td>
</tr>
<tr>
<td>Simultaneous neg. (C)</td>
<td>4.41</td>
<td>.85</td>
<td>.761</td>
<td></td>
</tr>
</tbody>
</table>

Treatment B: issue importance – single seller (high-medium-low) – multiple buyers (low-medium-high).
Treatment C: simultaneous negotiations – single seller – multiple buyers.
Treatment D: simultaneous negotiations – single seller – single buyer.
maintain the same levels of satisfaction and profit. If a buying team’s ordering of issues affects the impasse rate, we would expect to observe this effect when the most important issue to the buying team is bargained first (Treatment A). In contrast, bargainers fail to reach agreement when a buyer bargains the buying team’s least important issue first (Treatment B)—the scenario that should be the most conducive to the buyer reaching agreement.

In summary, choosing a simultaneous rather than a sequential agenda for bargaining with a buying team appears to have several results:

- higher profits for single sellers
- more integrative agreements
- greater buyer satisfaction.

Under sequential negotiations, issue order has a varied impact:

- bargaining issues in order of most-to-least rather than least-to-most important demonstrated no differences in buyers’ perceptions, behaviors or outcomes
- bargaining issues in order of most-to-least important raises the likelihood that a seller will break off negotiations.

4.2. Managerial implications

Conventional wisdom dictates that a single seller should avoid bargaining against multiple buyers due to the inherent imbalance in power and bargaining resources. However, salespeople need to reject this belief if they want to maximize the return on their bargaining efforts. Bargaining simultaneously with all members of a buying team can generate more integrative agreements than one-on-one bargaining, and a seller is unlikely to receive less profit than when bargaining all issues with a single buyer. Indeed, as long as the issues have the capacity for tradeoffs, there appears to be no benefit to bargaining one-on-one with a single buyer.

However, when a salesperson bargains with a buying team, the choice of a sequential or a simultaneous agenda is critical. Again, the choice appears to depend on the integrative nature of the issues to be bargained. If buyers and seller have different priorities regarding issues, possible tradeoffs could lead to integrative solutions and greater economic gain for the seller than the seller can achieve in sequential negotiations. In contrast, if buyers and a seller hold the same priorities regarding issues, a simultaneous agenda strategy would merely exacerbate the perceived power imbalance between the parties and place the seller at a disadvantage. In this case, tradeoffs would be difficult to affect and divide-the-pie solutions are likely to ensue (Jap, 1999). This finding does not mean that a salesperson must understand the exact importance of each of the issues to the parties. A salesperson only needs to understand that there are likely to be key differences in the importance of the issues to buyers and seller.

A simultaneous agenda strategy, with its mutually supportive environment and the achievement of greater profits for buyers, also leads to greater buyer satisfaction than a sequential agenda strategy does. Again, this finding is likely contrary to what a salesperson expects. A salesperson is likely to believe that his or her greater personal attention, which is possible during one-on-one interactions, is likely to promote greater satisfaction on the part of buyers. However, the diminished levels of buyer satisfaction found under sequential negotiations are likely to be detrimental to positive customer relationships.

Experiential and academic literature also suggests that there are benefits to bargaining issues in the order of least-to-most important when a sequential agenda is undertaken. A salesperson should understand that the reverse ordering of issues from most-to-least important does not automatically lead to lower profit. Rather, the critical concern is a greater probability that agreement will not be reached. This
This research was in part supported by Penn State’s Institute for the Study of Business Markets (ISBM).

Appendix A. Payoff matrices

<table>
<thead>
<tr>
<th>Seller profit tables ($ in millions)</th>
<th>Retailer margin</th>
<th>Advertising support</th>
<th>Credit terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option</td>
<td>Profit</td>
<td>Option</td>
<td>Profit</td>
</tr>
<tr>
<td>A</td>
<td>40</td>
<td>A</td>
<td>24</td>
</tr>
<tr>
<td>B</td>
<td>35</td>
<td>B</td>
<td>21</td>
</tr>
<tr>
<td>C</td>
<td>30</td>
<td>C</td>
<td>18</td>
</tr>
<tr>
<td>D</td>
<td>23</td>
<td>D</td>
<td>15</td>
</tr>
<tr>
<td>E</td>
<td>20</td>
<td>E</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>15</td>
<td>F</td>
<td>9</td>
</tr>
<tr>
<td>G</td>
<td>10</td>
<td>G</td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>H</td>
<td>3</td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>I</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Buyer profit tables ($ in millions)</th>
<th>Retailer margin</th>
<th>Advertising support</th>
<th>Credit terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Option</td>
<td>Profit</td>
<td>Option</td>
<td>Profit</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>A</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>B</td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>4</td>
<td>C</td>
<td>6</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>D</td>
<td>9</td>
</tr>
<tr>
<td>E</td>
<td>8</td>
<td>E</td>
<td>12</td>
</tr>
<tr>
<td>F</td>
<td>10</td>
<td>F</td>
<td>15</td>
</tr>
<tr>
<td>G</td>
<td>12</td>
<td>G</td>
<td>18</td>
</tr>
<tr>
<td>H</td>
<td>14</td>
<td>H</td>
<td>21</td>
</tr>
<tr>
<td>I</td>
<td>16</td>
<td>I</td>
<td>24</td>
</tr>
</tbody>
</table>

References


Acknowledgments

We thank the Editor, Area Editor and the reviewers for their constructive suggestions. We also thank Josh Eliasheff, Gary Lilien, Alison Lo, and Robert Wilken for their suggestions and comments on an earlier version of this article. We particularly appreciate Ethné Swartz, Maya Balakrishnan, and Wendy Cook for their generous and timely help. We are grateful to Jim Vona, Group Manager—Sales, Electronic Components Group, Panasonic Industrial Company; and the many anonymous salespeople and managers at a number of firms for their assistance. We especially acknowledge the outstanding research assistance of Philip Bertrand.

4.3. Limitations and directions for future research

Limitations of the main study include our employment of undergraduates as participants and the lack of explicit performance incentives. However, the subsequent use of MBA students to validate the study and the support provided by the practitioner survey provide credence to the findings. Nevertheless, additional replications and extensions in real world settings are warranted to assure the generalizability of the findings.

Other studies are needed to better explore the communications that take place between the buying team members when bargaining in simultaneous negotiations (Cooper & Kagel, 2005). Understanding these communications may illuminate how buyers within a group setting collaboratively interact to arrive at superior outcomes than buyers bargaining separately under sequential negotiations (Jap, Manolis, & Weitz, 1999). That is, it would be useful to peer into the “black box” of negotiations (Wilken, Cornelissen, Backhaus, & Schmitz, 2010). Within this communication context, a number of key factors could also be explored. Specifically, future studies could examine the manner in which intra-team member relationships affect bargaining processes and outcomes as well as the tenure and strength of relationships between buyers and sellers.

Our study also focused on three buyers concerned with three issues. An increase in the number of buyers representing the buying team and in the number of issues to be resolved would significantly increase the complexity of the bargaining task and the permutations involved in structuring the agenda. Accordingly, future research should examine a more diverse range of bargaining contexts by varying buying team size and composition; expanding the handling of each issue from a single buyer to a small group; bargaining a group of issues with multiple buyers and the rest of the issues on a one-on-one basis; or team selling approaches involving multiple sales personnel who interface with multiple buyers.

Determining how bargainers fare when specific negotiation styles are employed might provide valuable feedback. In this regard, gaining a better understanding of how training impacts negotiation processes and outcomes could also be revealing (Krishnamoorthy, Misra, & Prasad, 2005). Aspirations also warrant increased attention as recent research (Balakrishnan, Gomez, & Vohra, 2011) suggests that prior contractual arrangements may temper negotiators’ aspirations.

Finally, we are intrigued by the observations of Cateora, Gilly, and Graham (2011). They find that most business people in Western societies divide complex negotiation tasks into a series of smaller, sequential bargaining tasks; while business people in nonwestern societies tend to divide complex negotiation tasks into a series of smaller, sequential bargaining tasks. Moreover, these researchers suggest that negotiating team tactics may temper negotiators’ aspirations.

We thank the Editor, Area Editor and the reviewers for their constructive suggestions. We also thank Josh Eliasheff, Gary Lilien, Alison Lo, and Robert Wilken for their suggestions and comments on an earlier version of this article. We particularly appreciate Ethné Swartz, Maya Balakrishnan, and Wendy Cook for their generous and timely help. We are grateful to Jim Vona, Group Manager — Sales, Electronic Components Group, Panasonic Industrial Company; and the many anonymous salespeople and managers at a number of firms for their assistance. We especially acknowledge the outstanding research assistance of Philip Bertrand.
The multiple roles of interpersonal communication in new product growth

Trichy V. Krishnan a, P.B. “Seethu” Seetharaman b,⁎, Demetrios Vakratsas c

a NUS Business School, NUS, Singapore
b Washington University, St. Louis, USA
c McGill University, Canada

1. Introduction

The year 2010 began with the launch of a major new product by Apple: the iPad. While the iPad captured the world’s attention, it was only the 24th entrant in the e-book reader market! In fact, the very first e-book reader was the Librie, introduced by Sony in 2004–05. The Librie was followed by Sony’s Reader in 2006, Samsung’s Papyrus and Amazon’s Kindle in 2007 and so on (website ref: Wikipedia on e-book, 2012). By the beginning of 2012, forty different companies have introduced e-book readers, thirty of which feature electronic-paper displays (ex: Amazon Kindle) and ten of which feature non-electronic-paper displays (ex: Apple’s iPad, Kindle Fire). Interestingly, many of the major product launches occurred in the 2009–10 period, when the e-book reader market was still in its infancy (website ref: ebookreader.com, 2012). Why was that? Why couldn’t companies wait until the market reached the growth stage? One reason may be that the firms did not want to be late to exploit the snowballing/cascading effects of Interpersonal Communication (hereafter referred to as IPC for convenience), which includes, among other things, robust word-of-mouth effects and more recent phenomena such as the effects of social networking. IPC has been widely recognized for the important role it plays in the market adoption of a new product. From the perspective of a firm, the snowballing / cascading effects2 of IPC imply that, as the number of adopters of a brand increases, the rate of future adoptions of that brand also increases. The prospects of exponential growth in sales over time lure firms to launch their brands sooner rather than later, without waiting to see if the product category turns out to be successful or not.

In a recent paper that makes an important contribution to the marketing literature on IPC, Libai, Muller, and Peres (2009) show that IPC among previous adopters of a duopoly brand affect not only the brand’s own sales but also the sales of the competing brand. If one were to generalize this point to an oligopoly containing multiple (>2) brands, one would argue that IPC among previous adopters of a certain brand would influence product category sales as a whole (in addition to that brand’s own sales). The actual increase in product category sales resulting from a specific brand would depend on the attractiveness of that brand relative to other brands in the category. Therefore, a careful accounting of a brand’s IPC must separate its direct effect on the brand’s own sales from its indirect effect through the overall increase in product category sales. Despite the extensive literature on estimating IPC effects in new product growth models,3
a careful accounting of a brand’s IPC has not been undertaken. Addressing this gap in the literature is one of the purposes of this study.

It is widely acknowledged that IPC is not entirely within the manager’s control but happens, in large part, organically among the brand users (i.e., in an exogenous manner). What a brand manager typically does is manage the traditional marketing mix (i.e., price, advertising and distribution) for her brand. However, to optimize her marketing mix spending, the brand manager must still correctly understand and analyze the effects of the marketing mix on her brand’s sales relative to the effects of IPC (at both the brand-level and the category-level). Enabling this type of managerial decision-making is another purpose of this study.

The dual objectives of our study are summarized using the following two questions of interest to a brand manager:

1. Do previous adopters of a brand generate both category-level and brand-level IPC? If so, how do these two types of IPC work together in driving a brand’s sales?
2. How can we infer the impact of IPC on category sales and brand sales in the presence of marketing mix variables that affect new product growth?

Considering the recent rise in the significance of IPC marketing in the business sector (as evidenced by both the market success of books including “The Tipping Point” by Malcolm Gladwell and “The Secrets of IPC Marketing” by George Silverman and the advent of social networking websites including Facebook, Twitter and personal blogs), our study is both valuable and timely to brand managers operating in this environment.

Previous empirical studies focusing on new product growth (see Mahajan et al. (1990) and Peres et al. (2010) for reviews) typically measured IPC at the category level but neglected inter-brand competition and brand-level IPC. Our goal in this paper is to analyze new product growth and the effects of IPC therein while explicitly taking into account the effects of inter-brand competition and, in turn, brand-level IPC.

Next, we explain the difference between category-level and brand-level IPC.

2. Category-level versus brand-level IPC

IPC refers to the interpersonal communication between previous and potential adopters of a product. Such communication may occur directly (e.g., face-to-face, telephone, personal e-mail, personal observation of the product in use) or indirectly through social networks (e.g., online communities such as Facebook and Yahoo discussion groups, professional associations such as the AMA, viral videos on YouTube4) (Van den Bulte & Stremersch, 2004). Consumers perceive IPC as a more credible source of information about product quality/fit than company-sponsored communication such as advertising and sales promotions. By enabling potential adopters to learn about product attributes based on the user experiences of previous adopters, IPC helps to reduce consumer uncertainty about the potential benefits and weaknesses of a product. Traditional models of new product growth have assumed IPC effects to operate at the level of the product category only, ignoring IPC effects at the brand level (with the only exceptions being Krishnan, Bass, & Kumar, 2000; Libai et al., 2009; how our study differs from these studies will be explained in the next section).

To differentiate the effects of category-level IPC from the effects of brand-level IPC, we need to recognize two types of brand-level IPC effects. Let us explain this in more detail. Brands within a product category have both shared and unique features. For example, a large display screen and easy access to multimedia are common features that underlie all smart phones (in comparison to traditional mobile phones). However, single-touch access to YouTube is a unique feature that distinguishes the Apple iPhone from the BlackBerry5 and other smart phones. As users sample both the shared and unique features of the iPhone, their IPC will, in turn, involve both types of features. First, to the extent that a user’s IPC involves the efficacy of the common features, his or her brand-level IPC would influence category-level sales (i.e., through influencing the perceived attractiveness of smart phones in general). We call this the “brand-to-category IPC” effect. Second, to the extent that a user’s IPC involves the efficacy of the unique features of the iPhone, his or her brand-level IPC would influence brand-level sales (i.e., through influencing the perceived attractiveness of the iPhone in particular). We call this the “brand-to-brand IPC” effect.

In addition to the “brand-to-category IPC” discussed above, there is another type of category-level IPC that is generated from the category per se and not from any specific brands. This type of IPC involves information sources such as websites (e.g., Wikipedia) and internet-based social networks (e.g., Facebook), which focus on the product category as opposed to any particular brand. For example, a WSJ article published in 2010 (WSJ, 2010) described e-book readers as a revolutionary category challenging the older category of paper-based books. The Wikipedia website on e-books (website ref: Wikipedia on e-book, 2012) also discusses e-book readers in general as opposed to any specific e-book reader. In addition, yet another website (website ref: Parenthood.com, 2012) discusses how a family should decide between a minivan and an SUV, without referring to any specific brand in either category. This is the type of category-level IPC that is captured by traditional diffusion models such as the Bass model. We call this “category-to-category IPC.”

To summarize, we propose three distinct types of IPC. The first type is generated at the brand level but influences the adoption of the category (“brand-to-category IPC”). The second type is generated at the brand level and influences the adoption of that particular brand (“brand-to-brand IPC”). The third type is generated at the category level and influences the adoption of the category (“category-to-category IPC”). Next, we propose a category-cum-brand level growth model that simultaneously accommodates all three types of IPC effects and captures the effects of brand-level marketing variables. Our proposed model nests the Bass model of category-level diffusion as a special case (i.e., when only “category-to-category IPC” is present but “brand-to-category IPC” and “brand-to-brand IPC” are both absent). We estimate our proposed model using quarterly data on brand-level sales, prices, advertising expenditures and distribution in the SUV category.

Before presenting our model, we review the extant literature on the issue of accommodating and measuring different types of IPC effects on new product growth.

---

4 A good example of a successful viral marketing campaign is the “Will it Blend?” series of YouTube videos developed by BlendTec, in which George Wright—the VP of Marketing and Sales—was shown pulverizing iPhones, athletic shoes, marbles, Bic lighters etc. The 84 videos posted on YouTube were watched 200m times and increased sales by 700% for BlendTec blenders.

5 In their latter models (e.g., BlackBerry Torch), BlackBerry added this feature.

6 Another example pertaining to minivans is the following: By talking to current users of the Mercury Villager minivan, a potential adopter (a) experiences reduced uncertainty about the safety of minivans (as opposed to, say, sedans) and (b) learns more about the unique design features of the Mercury Villager compared to, say, the Nissan Quest. Whereas (a) represents “brand-to-category IPC”, (b) represents “brand-to-brand IPC.” Similarly, Consumer Reports publish reliability and owner satisfaction ratings on passenger cars and SUVs, which are based on surveys of existing users of particular brands. (http://www.consumerreports.org/cro/cars/new-cars/suv/index.htm). Although these reports pertain to different brands, they also provide information at the level of the category by organizing all its constituent brands in one common place. In this sense, Consumer Reports provides both “brand-to-brand IPC” and “brand-to-category IPC.”
3. Review of pertinent literature

Despite the vast marketing literature on new product diffusion, few papers have attempted to model and estimate the effects of IPC on brand-level sales growth. Two studies stand out as important exceptions: Krishnan et al. (2000) and Libai et al. (2009),7 henceforth referred to as KBK (2000) and LMP (2009), respectively. Below, we illustrate the differences between our modeling approach and the approaches used in the two previous studies.

In KBK (2000), category-level IPC is modeled to differentially affect the growth in sales of each brand over time. The objective of KBK (2000) is to propose a model that tracks diffusion at the brand level. Using data on cellular phone service adoptions in the US during the 1980s and 1990s, the authors report that the sales growth of each brand is affected by IPC from previous adopters of the category as a whole, regardless of the number of previous adopters adopting one brand versus another. Thus, the IPC modeled in KBK (2000) can be described as “category-to-brand IPC.” The authors ignore both types of brand-level IPC (i.e., “brand-to-brand IPC” and “brand-to-category IPC”). Additionally, the model treats category-to-category IPC as technically equal to the sum of all category-to-brand IPCs.

In LMP (2009), the sales growth of a brand in a duopoly market is allowed to depend on not only the IPC from its previous adopters but also the IPC from the previous adopters of the competing brand. Using data on cellular phone service adoptions in Europe during the 1980s and 1990s, the authors separately estimate the two components of brand-to-brand IPC on the sales growth of a given brand. This approach extends the category-to-brand IPC formulation of Krishnan et al. (2000) to further disentangle the separate contributions of the previous adopters of a brand from the previous adopters of the competing brand. The authors find that the competing-brand-to-focal-brand IPC component is roughly 55% of same-brand-to-brand IPC component. However, while focusing on brand-level IPC effects, the LMP model ignores category-level IPC.

In our study, we explicitly account for two types of category-level IPC effects (i.e., “brand-to-category” and “category-to-category”) in addition to brand-level IPC effects (“brand-to-brand”). In other words, our approach generalizes the models of KBK (2000) and LMP (2009) to provide a flexible representation of three types of IPC effects in new product markets with multiple brands. This is the first contribution of our study.

Neither KBK (2000) nor LMP (2009) explicitly models category-level adoption. Because the focus of both papers is on brand-level adoption, the only way to (indirectly) infer category-level adoptions is to sum the predicted adoptions across brands. However, such a modeling approach precludes the behavioral possibility that consumers may first decide whether to adopt a certain category before they decide which brand to adopt. Our modeling approach, in contrast, allows us to study whether a typical consumer in a market follows a two-stage (i.e., category in- first decision choice as a two-stage process. Thus, the second contribution of our study is to explicitly address this particular issue in our modeling framework. Nonetheless, we must first note that, given our use of an aggregate-level model, our inferences are necessarily at the overall market level rather than the individual consumer level.

Last but not the least, the brand-level diffusion models of KBK (2000) and LMP (2009) are wieldy only for product categories containing two or three brands and given that brands enter the market simultaneously. For categories involving three or more brands or the sequential entry of different brands, it becomes non-trivial to modify and estimate the KBK (2000) model. Because the LMP model addresses the case of two brands, it can easily accommodate own-brand and cross-brand effects. However, extending the LMP model to include three or more brands would necessitate the estimation of too many cross-brand parameters and render the model unwieldy. In contrast, our proposed model estimates both within-brand and category-level IPC effects, enabling us to capture the residual category-level IPC effects that remain after accounting for within-brand IPC effects. In this manner, our brand-level diffusion modeling approach is designed to easily handle a large number of brands as well as multiple entries and exits of brands over time. This is the third contribution of our study.

Taken together, the above discussion can be summarized as follows. First, our proposed model captures the effects of three different types of IPC (i.e., brand-to-brand, brand-to-category and category-to-category) on new product growth. Second, our proposed model allows us to test whether the market reflects a one-stage or two-stage consumer adoption process. Third, as a brand-level diffusion model, our proposed model is suitable for use with product categories that contain a large number of brands and involve sequential entries and exits of brands.

In addition, we measure the effects of marketing variables on brand-level sales growth, which is ignored in KBK (2000). This consideration is important because a brand manager needs to understand the relative marginal impact of IPC versus marketing mix effects in driving brand-level sales growth.

Before we close this section, we need to highlight some key differences between a model recently proposed by Landsman and Givon (2010), hereafter referred to as LG, and our proposed model. The LG article offers an individual-level statistical model to describe the process by which a customer adopts a new service program offered by a commercial bank. The customer is modeled as moving from the consideration stage to the brand choice stage while choosing among various plans offered in the new service program. The authors employ the proportional hazard model for the consideration stage and the multinomial logit model for the brand choice stage. An important aspect of the LG empirical application is that, in each period, the individual is allowed to switch plans or even dis-adopt (i.e., quit) the new service program altogether. This aspect of the model means that the risk associated with the individual’s original adoption decision is low. However, with regard to a new durable product (such as an SUV or an iPad), the risk and uncertainty associated with the original adoption is much higher because neither disadoption (i.e., product return) nor period-to-period switching is likely to take place. In fact, as the authors suggest in their article, “One could try to develop [such] a model from overall utility maximization principles.” Our model takes LG up on this suggestion.

4. Statistical model of IPC in new product growth

Below we present a statistical model of category-level and brand-level sales growth that accommodates (1) all three IPC effects (brand-to-brand, brand-to-category and category-to-category) and (2) the effects of marketing mix variables. The proposed model is utility-theoretic (the model derivation from utility maximization primitives is provided in the Appendix) and provides an explicit empirical test of whether the market adoption of the product is a two-stage (i.e., category adoption followed by

---

7 Other studies that have estimated brand-level new product diffusion models include Peterson and Mahajan (1978), Parker and Gattigsson (1994), Shankar, Carpenter, and Krishnamurthi (1998), Kim, Bridges, and Sivriyazivatva (1999) and Jun, Kim, Park, Park, & Wilson (2002). However, these studies do not specifically focus on brand-level IPC. We also do not consider brand-choice models, such as that proposed by Mahajan, Sharma, and Buzzell (1993), in which it is assumed that the adoption of a new brand is affected only by its own previous adopters.
brand choice) or one-stage (simultaneous choice of category adoption and brand choice) decision, as explained earlier.

The sales of brand $j$ at time $t$ are given by the following:

$$Sales_j = [M-CS(t)] + \sum_{t-1}^{\infty} e^{\beta t} \cdot \frac{e^{\beta t}}{1 + e^{\beta t}} \cdot \sum_{k=1}^{\infty} e^{\lambda t}$$

(1)

where $M$ stands for the market potential (i.e., total unit sales in the product category over its lifetime), $CS(t)$ stands for the cumulative product category sales prior to time $t$ and $z_t$ stands for the product category attractiveness at time $t$ and is given by the following:

$$z_t = \ln\left(\ln\left(\frac{1-F_{t-1}}{1-F_t}\right)\right) + \gamma \ln \sum_{k=1}^{\infty} e^{\lambda t} + \delta \cdot XC_t$$

(2)

where $F_t$ stands for the cumulative distribution function, evaluated at time $t$, characterizing the baseline hazard process of consumers’ product adoption times (assumed to be the Bass Model), $XC_t$ stands for external economic factors that affect the whole category. For our study, we use the retail price of gasoline at time $t$. $V_{ij}$ stands for the attractiveness of brand $j$ at time $t$ and is given by the following:

$$V_{ij} = \alpha_0 + \beta_1 \cdot Price_i + \beta_2 \cdot AdStock_j + \beta_3 \cdot Dist_j + \beta_4 \cdot CumSales_{ij,t-1} + \beta_5 \cdot OE_i$$

(3)

where $Price_i$, $AdStock_j$, and $Dist_j$ stand respectively for the price, advertising-stock and distribution associated with brand $j$ at time $t$. $CumSales_{ij,t-1}$ stands for the cumulative sales associated with brand $j$ up to $t-1$ and $OE_i$ stands for the order of entry position for brand $j$. The fifth term on the RHS of Eq. (3), i.e., $\beta_4 \cdot CumSales_{ij,t-1}$, represents the effects of brand-level IPC generated by previous adopters of brand $j$. The parameter $\beta_4$ captures the impact of this IPC on brand choice, which we call brand-to-brand IPC. As mentioned earlier, the motivation for this term comes from LMP (2009), where it is shown that the sales growth of a brand depends on brand-level IPC. Advertising stock is defined as follows:

$$AdStock_j = Adv_j + \lambda \cdot Adv_{j-1} + \lambda^2 \cdot Adv_{j-2} + \ldots$$

(4)

where $0 \leq \lambda \leq 1$ is the advertising carryover parameter. This is the familiar ‘Koyck lag formulation’ that has been extensively used to model the dynamic effects of advertising on demand.

Let us explain the model as represented in Eq. (1). The first two terms on the RHS of Eq. (1), i.e., $[M-CS(t)]$ and $e^{\beta t} / (1 + e^{\beta t})$, together represent product category sales at time $t$, whereas the third term, i.e., $e^{\lambda t} / \Sigma e^{\lambda t}$, represents the conditional (on current product category sales) market share for brand $j$ at time $t$. Eq. (1), therefore, decomposes brand-level sales into two multiplicative components: (1) product category sales and (2) brand market share.

Next consider Eq. (2), which represents the product category attractiveness at time $t$. The RHS of Eq. (2) contains three terms. We explain each term as follows.

1) The first term, $ln[n ln(1-F_{t-1})]$, represents the effect of the baseline hazard process that characterizes consumers’ product adoption times. The reason for choosing this particular $ln[ln()]$ function is explained in the Appendix. In short, we show that if $F_t$ follows the cumulative distribution function associated with the Bass model, our product category sales model reduces to the Bass model with the underlying parameters $p$ and $q$ representing the innovation and imitation coefficients of the Bass model. The parameter $q$ specifically represents the impact of category-to-category IPC. If $q$ equals 0, our product category sales model reduces to the Geometric model, which assumes a constant [i.e., memory-less] baseline hazard.

2) Separate from its ability to reduce to the Bass model, let us provide an intuitive explanation of what the function $ln[ln(1-F_{t-1})] / (1-F_t)$ represents. Let us first consider $\frac{\frac{\gamma \cdot \lambda}{p^2 + \lambda^2}}{1 + e^{\beta t}}$. The integrand on the RHS is the hazard function, i.e., conditional probability of adoption at time $t$ given that adoption has not yet occurred. The integral, which represents the integrated hazard over the interval representing discrete period $t$, represents the hazard in discrete time. Taking another logarithm yields the log integrated hazard, which, when exponentiated (as would be done in the random utility formulation with Gumbel errors), yields the discrete hazard representing the consumer’s baseline probability of purchasing over the discrete period $t$.

3) The second term in Eq. (2), i.e., $\ln \Sigma e^{\lambda t}$, also called inclusive value, represents the collective attractiveness of all brand-level activities and the coefficient $\gamma$ represents its effect on the category-level adoption at time $t$ (Campo, Gjibscrechts, Goossens, & Verhetsel, 2000). Brand-level activities include the marketing mix elements employed by the brands (i.e., advertising, pricing and distribution) and the brand-level IPC generated by the previous adopters of the brands. How do these brand-level activities influence the attitude of consumers toward category adoption? Consider, for example, an advertising message carried out by a company to create awareness of their brand in the marketplace. Such an advertising message will also increase the awareness of the category in general, which will, in turn, influence category adoption decision. Alternatively, if we consider the type of advertising that is designed to convey the value of a specific brand (i.e., brand’s USP) to consumers, we might not expect a strong influence on category adoption. Nonetheless, some effect is still expected because the value proposition of the brand signals to some extent the value of the encompassing category. Thus, regardless of the type of advertising carried out by a brand, there will always be some effect on category-level adoption decisions. Similarly, the prices of the various brands to which consumers are exposed collectively influence how consumers judge the value of the whole category with respect to the prevailing prices. Additionally, the IPC generated by the adopters of the various brands will collectively reduce the uncertainty that a consumer may be facing at the category level, especially those points of uncertainty that are common across all the brands. In essence, we can expect brand-level activities to collectively and significantly influence a consumer’s category-level adoption.

In the proposed model, this collective force from brand-level activities is represented by $\ln \Sigma e^{\lambda t}$ and its impact on category-level sales is captured by the parameter $\gamma$, also called the inclusive value parameter. Note that within the collective force, individual parameters signify the relative contribution from each effort type, namely, $\beta_1$ for Price, $\beta_2$ for advertising, $\beta_3$ for distribution and $\beta_4$ for IPC. Hence, the impact of brand-level IPC on category adoption, which we call brand-to-category IPC, is captured through the $\gamma, \beta_4$ parameter set.

4) Given that we have two factors influencing category-level sales, namely, baseline hazard (first term of Eq. (2)) and the collective force from brand-level activities (second term of Eq. (2)), we will now explain how they work together to influence category-level sales. Summing Eq. (1) over all the brands, we obtain the category-level sales function.

$$Sales_c = [M-CS(t)] \cdot \frac{e^{\beta t}}{1 + e^{\beta t}}$$

(5)

6 Considering that $\beta_4$, taken alone, represents brand-to-brand IPC, we notice that in our model formulation, brand-to-category IPC will not exist unless brand-to-brand IPC exists. This restriction occurs because we derive our model from utility maximization primitives (using the nested logit formulation). Relaxing this assumption, beyond violating utility maximization, would entail estimating a separate brand-to-category IPC effect for each brand, which we avoid for the sake of model parsimony.
where  
\[ z_t = \ln \left( \frac{1 - F_t}{1 - F_{t-1}} \right) + \gamma \ln \left( \sum_{i=1}^k \rho_i \right) + \delta \times \text{GasPrice}_t, \]

the \( k \) subscript in the second term refers to brand \( k \). Including the full expansion of \( z_t \) in the sales function and ignoring the \( \text{GasPrice}_t \) term for the sake of expositional clarity, we obtain the following function for category-level sales.

\[ \text{Sales}_t = |M - \text{CS}(t)| \times \left[ \ln \left( \frac{1 - F_t}{1 - F_{t-1}} \right) \right] \left( \sum_{i=1}^k \rho_i \right) \left\{ \left( \sum_{i=1}^k \rho_i \right) \right\} \left\{ \left( \sum_{i=1}^k \rho_i \right) \right\} \]

For further expositional clarity, consider only the advertising term within \( \text{V}_{\text{Ad}} \), i.e., \( \beta_2 \times \text{AdStock}_t \). Thus, the category-level sales function can be re-written as follows:

\[ \text{Sales}_t = |M - \text{CS}(t)| \times \left[ \ln \left( \frac{1 - F_t}{1 - F_{t-1}} \right) \right] \left( \sum_{i=1}^k \rho_i \right) \left\{ \left( \sum_{i=1}^k \rho_i \right) \right\}, \]  

(5)

As explained earlier in Point 2, note that \( \ln \left( \frac{1 - F_t}{1 - F_{t-1}} \right) \) represents the conditional probability of adoption happening in the interval \((t-1, t]\). Consider the case of \( \gamma = 0 \). Eq. (5) reduces to the baseline category-level adoption model i.e., the Bass model (see Appendix for the derivation) because the advertising efforts by various brands will have no effect on category-level diffusion, regardless of the value of \( \beta_2 \). This statement implies that, if \( \gamma \) is significantly greater than 0, then brand-level advertising will collectively impact category-level sales by enhancing the baseline category adoption probability through the parameter set \( \{\gamma, \beta_2\} \). Note that the baseline category adoption probability, as represented by the Bass model, is characterized by two parameters \( q \) and \( p \). The former represents the impact of category-to-category IPC and the latter represents the extent to which customers acutely need the product or their ability to analyze the product and decide for themselves without relying on category-to-category IPC. Thus, the brand-level activities (the second term of Eq. (2)) enhance the baseline hazard (the first term of Eq. (2)) and, in turn, influence the category-level sales.

5) The parameter \( \gamma \) also tells us if the market-level adoption follows a two-stage (in which case \( \gamma \) would be significantly less than 1 but positive) or one-stage (in which case \( \gamma \) will be close to 1) decision process (see Ben-Akiva & Lerman, 1985). If \( \gamma \) turns out to approach zero, then the category incidence decision and brand choice decision are independent, implying that the two are separate decisions measured at the market level.

6) The last term, \( \delta \times \text{GasPrice}_t \), represents the effect of quarterly gasoline prices on the demand for SUVs, with \( \delta \) expected to be negative.

To summarize, our model proposes three distinct roles for IPC in the sales growth of new product categories: the brand-to-brand IPC effect as captured by the parameter \( \beta_3 \), the brand-to-category IPC effect as captured jointly by the parameter set \( \{\gamma, \beta_2\} \) and the category-to-category IPC effect as captured by the parameter \( q \). Whereas previous research on the sales growth of new products focused on only the first role of IPC, the present study is the first to account for the second and third roles of IPC. Our model achieves this by simultaneously accommodating the effects of marketing variables on brand-level sales growth. Another attractive feature of our proposed model is that it nests the popular Bass model as a special case (when brand-to-brand IPC effects, brand-to-category IPC effects and brand-level marketing mix effects are all absent). The various IPC types are depicted in Fig. 1.

However, one should understand that it is not possible to draw conclusions about the relative impact of the three IPC types by simply looking at the respective parameter estimates. There are three reasons. First, the proposed model is highly non-linear. Second, IPC effects are dynamic and recursive by nature. Third, brand-to-category IPC shares the \( \beta_3 \) with brand-to-brand IPC and shares the \( \gamma \) with the marketing mix variables. Hence, it is necessary to run the simulation exercise to analyze comprehensively the differential impact of the three types of IPC on the sales of a brand. We will demonstrate this simulation exercise in a later section.

5. Empirical results

We use sales data from the SUV (Sport Utility Vehicle) market. Let us first offer a brief description of this category and the market.

5.1. Market and data description

The SUV is considered as a type of cross-over design between a regular car and a pick-up truck. Cross-over designs of cars began appearing in the 1960s and the evolution toward the truck-cum-car began in the 1970s (website ref: AutoTrader.com, 2007). The SUV was born in the 1980s. The early SUVs had a large, rugged, off-road look with the passenger/cargo space ratio favoring cargo space and were used predominantly in rural towns and less in suburbs and urban cities. As the SUV design further evolved, these large SUVs gradually vanished from the market. Quarterly sales of the large SUVs, which were approximately 65,000 in the 1980s, dropped to approximately 40,000 by 1992 and further dropped to an insignificant amount by the end of 1993. The ‘standard’ SUVs, which co-existed with their larger counterparts in the 1980s, evolved to include more creature comforts and higher ratios of passenger/cargo space. The quarterly sales of these standard SUVs, however, hovered at approximately 70,000 for 4 years (1986–1990) before eventually taking off in the year 1990. Although one could take any quarter in the late 1980s as the starting period for the analysis, we select the third quarter of 1990 for two reasons. First, as shown in Fig. 2, the sales curve of the SUV category flattens out until 1990, with a recovery/rise beginning in the third quarter of 1990. This pattern signifies the start of a new cyclical phase. We later explicitly check the sensitivity of our empirical estimates to this starting point.

Second, families-with-young-babies who adopted minivans in the 1980s became families-with-teenagers who embraced SUV adoption in the 1990s. Accordingly, in the 1990s, SUV designs began to emphasize road capabilities and in-vehicle comfort, which made SUVs even more appealing to urban and suburban households (as opposed to the SUVs of the 1980s that primarily appealed to rural households). Hence, in our study, we focus our attention to the category of small, suburban SUVs that were introduced in the 1990s, with the analysis period of our data set spanning 42 quarters, from the third quarter of 1990 to the last quarter of 2000. Our data include quarterly brand-level sales, advertising expenditures, list prices and distribution reach in terms of number of dealers. Table 1 reports the average market shares of available brands over the study period.

9] In fact, because of this demographic shift in the primary target market in the 1990s, companies reformulated their SUVs (e.g., Chevy Blazer, GMC Jimmy) to feature creature comforts and form (e.g., Chevy Blazer-S, GMC Jimmy-S). The resulting suburban SUVs of the 1990s had much broader appeal than their ancestors. As a result, sales of SUVs rapidly ramped up in the 1990s, with sales in that decade overshadowing sales in the previous decade by a factor of 4.

10] While list prices and distribution reach remained unchanged over all quarters within a given year, sales and advertising expenditures vary from quarter to quarter. Sales and list prices were obtained from Ward’s Automotive Yearbooks (1990–2000). Distribution data were obtained from Automotive News. Data on advertising expenditure were obtained from CMR/TNS (Competitive Media Reporting), which included information on brand expenditures in ten major types of media.
5.2. Refinements of data structure and model variables

Before estimating our proposed model, we undertake the following refinements.

1. Structural break: The luxury SUVs started appearing on the market in January–March 1995, which corresponds to Quarter 19 in our data. Although this marked an important strategic event in the SUV category, when we look at the category-level diffusion curve (see Fig. 2), we hardly discern any significant jump or break in category sales (that is inclusive of luxury SUV sales) occurring around quarter 19 in year 1995. This trend suggests that, rather than drawing an entirely new pool of customers to the market, luxury SUVs catered to a similar pool of customers as standard SUVs. An obvious question that arises pertains to whether and how existing SUV brands were affected by and responded to the entry of luxury SUVs. It is conceivable that, on account of this “structural break” in Quarter 19, (a) some existing brands repositioned themselves and altered their marketing mix strategies, (b) marketing strategies of brands began having different impact on sales and (c) IPC began having different impacts on sales.11 To accommodate such changes in the marketplace due to the structural break, we extend our proposed model in four ways.

First, we introduce additional brand intercepts to represent the potential repositioning of existing brands after Quarter 19. Second, we allow for two different sets of marketing mix coefficients after Quarter 19: one for standard SUVs and another for luxury SUVs both of which are distinct from the marketing mix coefficients prior to Quarter 19. This division results in three sets of marketing mix coefficients, one for each of the three “regimes,” i.e., [Regime 1], [Regime 2L], and [Regime 2S], where “L” and “S” refer, respectively to “luxury” and “standard” SUVs in the post-Quarter 19 period. When calculating the impact of the Bass baseline hazard in [Regime 2], however, we use the combined sales of SUVs in the first 18 quarters as the starting value of lagged cumulative SUV sales in Quarter 19.

Third, we allow for the possibility for brand-to-brand IPC effects in [Regime 2L] and [Regime 2S] to differ from the effects in [Regime 1]. For example, consumers spread information about specific brands more or less by word of mouth when new sub-categories are formed within the SUV category. One reason is that consumers are now more

---

11 It is interesting to note that even after the introduction of the luxury SUVs, standard SUVs continued to co-exist and sell well. This is in marked contrast to the situation in 1990 when ‘large SUVs’ went off the market with the arrival of the suburban SUVs.
positively drawn to brands, which increases the positive impact of IPC. An alternate possibility is that consumers are starting to pay more attention to advertising efforts of different brands to better understand the differences between brands, which would lower the net impact of brand-to-brand IPC. To accommodate such shifts in IPC effects, we include cumulative brand sales from Quarter 19 as an additional covariate in the consumer’s brand attractiveness equations in [Regime 2L] and [Regime 2S]. The following three brand attractiveness equations for [Regime 1], [Regime 2L] and [Regime 2S], respectively, summarize the three model extensions discussed above.

\[ V_{jt} = \alpha_{jt} + \beta_{jt1} \cdot \text{Price}_{jt} + \beta_{jt2} \cdot \text{AdStock}_{jt} + \beta_{jt3} \cdot \text{Dist}_{jt} + \beta_{jt4} \cdot \text{CumSales}_{jt-1} + \beta_{jt5} \cdot \text{OE}_{jt}. \]  

(6)

\[ V_{jt} = \alpha_{jt} + \beta_{jt1} \cdot \text{Price}_{jt} + \beta_{jt2} \cdot \text{AdStock}_{jt} + \beta_{jt3} \cdot \text{Dist}_{jt} + \beta_{jt4} \cdot \text{CumSales}_{jt-1} + \beta_{jt5} \cdot \text{CumSalesFromQ19}_{jt} + \beta_{jt6} \cdot \text{OE}_{jt}. \]  

(7)

\[ V_{jt} = \alpha_{jt} + \beta_{jt1} \cdot \text{Price}_{jt} + \beta_{jt2} \cdot \text{AdStock}_{jt} + \beta_{jt3} \cdot \text{Dist}_{jt} + \beta_{jt4} \cdot \text{CumSales}_{jt-1} + \beta_{jt5} \cdot \text{CumSalesFromQ19}_{jt} + \beta_{jt6} \cdot \text{OE}_{jt}. \]  

(8)

where \( \beta_{jt2} \) and \( \beta_{jt5} \) are the parameters that capture the shift in IPC that might have happened after the entry of the luxury brands.

Fourth, we allow for the possibility that category-to-category IPC may change after Quarter 19. To capture this, we allow the word-of-mouth parameter in the Bass Model, \( \eta \), to take a different value starting from Quarter 19. Correspondingly, the first term in Eq. (2) reflects a different baseline hazard process starting from Quarter 19.

2. Replacement purchases: The proposed model pertains to first-time purchases only. Because our study period covers 10 years, it is possible for the observed sales in later years to include both first-time purchases and replacement purchases. To account for this possibility, we first impute the number of replacement purchases in each quarter. Next, we subtract these imputed replacement purchases from the observed sales in each quarter to obtain a time series of first-time purchases. Finally, we use the time series of derived first-time purchases to fit the proposed model (details about the imputation procedure are provided in the Appendix). 12

3. Order of entry: The direct effect of brand-level IPC is accounted for in our model by including brand-level cumulative sales as a covariate in the brand attractiveness equation (see Eq. (3)). This variable captures, in part, differences among brands in their order of entry because the covariate is likely to take large values over the latter part of the study period for early entrants but only small values for later entrants. To remove such “confounding cross-brand effects” of order of entry and to explicitly study the effects of order of entry on brand sales growth, we include order of entry as a covariate in the brand attractiveness equation. Doing so allows the estimated coefficient of brand-level cumulative sales (\( \beta_{jt4} \)) to represent more clearly the impact of brand-to-brand IPC.

12 Instead of the iterative estimation procedure that we adopt in this study, one can simply use the average replacement cycle of SUVs, as reported in trade journals, to remove replacement purchases from the data. However, such a “deterministic” imputation ignores the fact that replacement purchases may be characterized by a stochastic diffusion process in addition to being influenced by IPC and marketing activities of brands. Our imputation approach accommodates such a consumer decision process for replacement purchases. For the sake of computational convenience, however, we assume that the same set of model parameters characterizes both first-time and repeat purchases. Relaxing this assumption and estimating a different set of parameters for replacement purchases would substantially increase the computational burden associated with estimation for reasons that are not germane to the issues under study in this paper.

4. Top 4 brands: As shown in Table 1, it is evident that 4 brands (Ford Explorer, Jeep Grand Cherokee, Chevy Blazer-S and Jeep Cherokee) dominate the SUV market during our study period, with a collective market share of 60%. Therefore, we estimate brand intercepts for only these 4 brands, setting the remaining brand intercepts to zero. However, in the estimation, we consider every brand with its own sales, price, advertising stock and distribution numbers. As a result, all of the estimated parameters are based on the sales growth evolution of all the 15 brands in the market. Only the intercept term is assumed to be zero for all the brands except the top four.

Table 1

<table>
<thead>
<tr>
<th>Brand</th>
<th>Manufacturer</th>
<th>Average market share</th>
<th>Time of entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ford Explorer</td>
<td>Ford</td>
<td>24.3%</td>
<td>Q3 1990</td>
</tr>
<tr>
<td>Jeep Grand Cherokee</td>
<td>Chrysler</td>
<td>15.3%</td>
<td>Q1 1992</td>
</tr>
<tr>
<td>Chevrolet Blazer-S</td>
<td>GM</td>
<td>13.9%</td>
<td>Q3 1982</td>
</tr>
<tr>
<td>Jeep Cherokee</td>
<td>Chrysler</td>
<td>10.0%</td>
<td>Q3 1974</td>
</tr>
<tr>
<td>Honda CRV</td>
<td>Honda</td>
<td>5.2%</td>
<td>Q1 1997</td>
</tr>
<tr>
<td>Jeep Wrangler</td>
<td>Chrysler</td>
<td>5.0%</td>
<td>Q1 1986</td>
</tr>
<tr>
<td>Toyota 4-Runner</td>
<td>Toyota</td>
<td>4.8%</td>
<td>Q2 1984</td>
</tr>
<tr>
<td>GMC Jimmy-S</td>
<td>GM</td>
<td>4.2%</td>
<td>Q4 1982</td>
</tr>
<tr>
<td>Nissan Pathfinder</td>
<td>Nissan</td>
<td>4.1%</td>
<td>Q4 1986</td>
</tr>
<tr>
<td>Nissan Xterra</td>
<td>Nissan</td>
<td>4%</td>
<td>Q2 1999</td>
</tr>
<tr>
<td>Isuzu Rodeo</td>
<td>Isuzu</td>
<td>3.6%</td>
<td>Q2 1990</td>
</tr>
<tr>
<td>Chevrolet Tracker</td>
<td>GM</td>
<td>2.8%</td>
<td>Q3 1988</td>
</tr>
<tr>
<td>Mercury Mountaineer</td>
<td>Ford</td>
<td>2.4%</td>
<td>Q2 1996</td>
</tr>
<tr>
<td>Subaru Forester</td>
<td>Subaru</td>
<td>2.2%</td>
<td>Q2 1997</td>
</tr>
<tr>
<td>Mazda Tribute</td>
<td>Mazda</td>
<td>2.1%</td>
<td>Q3 2000</td>
</tr>
<tr>
<td>Isuzu Trooper</td>
<td>Isuzu</td>
<td>1.8%</td>
<td>Q1 1984</td>
</tr>
<tr>
<td>Kia Sportage</td>
<td>Kia</td>
<td>1.5%</td>
<td>Q1 1995</td>
</tr>
<tr>
<td>Suzuki Vitara</td>
<td>Suzuki</td>
<td>1.5%</td>
<td>Q3 1998</td>
</tr>
<tr>
<td>Honda Passport</td>
<td>Honda</td>
<td>1.4%</td>
<td>Q4 1993</td>
</tr>
<tr>
<td>Mitsubishi Montero</td>
<td>Mitsubishi</td>
<td>1.3%</td>
<td>Q4 1996</td>
</tr>
<tr>
<td>Oldsmobile Bravada</td>
<td>GM</td>
<td>1.2%</td>
<td>Q4 1990</td>
</tr>
<tr>
<td>Suzuki Sidekick</td>
<td>Suzuki</td>
<td>1.1%</td>
<td>Q4 1988</td>
</tr>
<tr>
<td>Pontiac Aztek</td>
<td>GM</td>
<td>1.1%</td>
<td>Q3 2000</td>
</tr>
<tr>
<td>Hyundai Santa Fe</td>
<td>Hyundai</td>
<td>1.1%</td>
<td>Q3 2000</td>
</tr>
<tr>
<td>Isuzu Amigo</td>
<td>Isuzu</td>
<td>0.5%</td>
<td>Q1 1989</td>
</tr>
<tr>
<td>Suzuki X-90</td>
<td>Suzuki</td>
<td>-0.1%</td>
<td>Q1 1995</td>
</tr>
<tr>
<td>Isuzu Vehicross</td>
<td>Isuzu</td>
<td>-0.1%</td>
<td>Q1 1999</td>
</tr>
</tbody>
</table>

\[ LL = -84,380,232; \] (Category incidence \( LL = -56,873,605 \), Brand choice \( LL = -27,957,628 \), BCE = -169,660,970. All estimates are statistically significant at the 0.01 level. M was assumed to be 18,318,797. However, changing M by 20% did not significantly change the estimates. The advertising carry-over coefficient was identified through a grid search (based on the maximum likelihood criterion).

Table 2

Estimation results for proposed model.

<table>
<thead>
<tr>
<th>Period</th>
<th>Regime 1</th>
<th>Regime 2S</th>
<th>Regime 2L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category-level parameters</td>
<td>( \gamma )</td>
<td>0.3372</td>
<td>0.3372</td>
</tr>
<tr>
<td></td>
<td>( p )</td>
<td>0.0047</td>
<td>0.0047</td>
</tr>
<tr>
<td></td>
<td>( q )</td>
<td>0.0655</td>
<td>0.0687</td>
</tr>
<tr>
<td></td>
<td>Gas price</td>
<td>-0.0018</td>
<td>-0.0018</td>
</tr>
<tr>
<td>Brand-level parameters</td>
<td>( \alpha_{Ford Explorer} )</td>
<td>2.0654</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>( \alpha_{Chevy Blazer} )</td>
<td>0.8886</td>
<td>1.0925</td>
</tr>
<tr>
<td></td>
<td>( \alpha_{Jeep Cherokee} )</td>
<td>0.6905</td>
<td>0.2178</td>
</tr>
<tr>
<td></td>
<td>( \alpha_{Grand Cherokee} )</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td></td>
<td>Price</td>
<td>-0.8966</td>
<td>-0.8080</td>
</tr>
<tr>
<td></td>
<td>Advertising</td>
<td>0.3964</td>
<td>0.1610</td>
</tr>
<tr>
<td></td>
<td>Distribution</td>
<td>0.1949</td>
<td>-1.3392</td>
</tr>
<tr>
<td></td>
<td>Lag cum sales</td>
<td>3.8854</td>
<td>3.8854 and</td>
</tr>
<tr>
<td></td>
<td>Order of entry</td>
<td>-0.0707</td>
<td>-0.0707</td>
</tr>
<tr>
<td></td>
<td>Adv carryover</td>
<td>0.21</td>
<td>0.21</td>
</tr>
</tbody>
</table>

\[ a \] Refers to the time of introduction of the original model.
5.3. Estimation results and key findings

We use maximum likelihood to estimate model parameters (see Appendix for details). We present the estimation results in Table 2. Key findings are summarized below.

1. The estimate of the parameter representing the impact of brand-to-brand IPC, \( \beta_3 \), is positive and significant. This result implies that the direct effects of brand-level IPC on brand-level sales growth are significant.

2. The inclusive value coefficient (\( \gamma \)), representing the collective impact of brand-level activities (that includes brand-level IPC) on category sales, is positive and significant. This property, together with the finding that \( \beta_4 \) is significant, implies that the indirect effects of brand-level IPC on category-level sales growth are significant. Thus, our study presents the first documentation of this particular type of IPC in the marketing literature.

3. The estimated value of \( \gamma \) is 0.3372 and the estimated standard error is 0.0064. Therefore, the estimate of \( \gamma \) is significantly different both from 1 and from 0. This finding implies that the adoption process of the typical consumer in the SUV category follows a two-stage process, i.e., category incidence followed by brand choice. This finding simultaneously implies that the adoption process is neither a single-stage process in which category incidence and brand choice are decided simultaneously nor one in which the two decisions are independent of each other. Behaviorally, the finding suggests that a typical consumer decides whether to buy an SUV before deciding on the specific brand of SUV. The finding, in fact, corroborates the assumption made in LG (2010). Note, however, that our finding regarding the two-stage decision process pertains to a typical consumer (i.e., a representative consumer) given the use of aggregate-level sales data in the estimation. It is possible for the market to contain only consumers who undertake one-stage decision process and consumers who treat the two decisions as independent but that, due to the ‘averaging’ effect, we obtain an estimate suggesting that all consumers make the two-stage decision process. In other words, in the SUV market, there may be heterogeneity across consumers in the decision process, with some consumers undertaking a two-stage decision process, some undertaking a one-stage decision process and others treating the two decisions as independent. In this heterogeneous case, SUV manufacturing firms must identify the different consumer segments and target each segment with different marketing strategies. This course of action would, of course, require data on individual-level adoption (as in Landsman & Givon, 2010).

4. Findings 2 and 3 above imply that marketing mix variables indirectly influence category-level sales growth through the inclusive value term (i.e., the second term on the RHS of Eq. (2)) and their corresponding \( \beta \) values. Category-level diffusion models, such as the Generalized Bass Model (Bass, Krishnan, & Jain, 1994) typically impute composite measures of marketing variables at the category-level by aggregating across brands (using subjectively chosen weights, e.g., market shares). Identifying the consequences of such aggregation, in terms of possibly biasing category-level sales forecasts, is an issue worthy of future research.

5. The estimated category-to-category IPC parameter (\( q \)) is positive and significant both before and after the structural break (Quarter 19), which implies that category-to-category IPC effects are significant even after accounting for brand-to-category IPC effects (as discussed in findings 1 and 2 above). This finding suggests that at least some aspects of consumer IPC in an evolving product category pertain to the product category as a whole as opposed to specific brands in particular.

6. The estimated value of the category-to-category IPC parameter (\( q \)) is similar between \( \text{Regime 1} (0.0655) \) and \( \text{Regime 2} (0.0687) \). This finding, which is consistent with the smooth category diffusion pattern in Fig. 2 that shows no discernible jump in category sales after Quarter 19, suggests that consumers’ baseline adoption decisions for the SUV category have not changed significantly in terms of the parameters underlying the baseline hazard even after the introduction of luxury SUVs.

7. The estimated category-level innovation coefficient (\( \gamma \)) is positive and significant (0.0047). This coefficient captures the extent of baseline adoption by innovators, absent IPC and sensitivity to the marketing mix, which characterizes the SUV category. It reflects the level of intrinsic transportation needs of innovative consumers that are satisfied by SUVs.

8. The estimated marketing mix coefficients (\( \beta_3, \beta_4, \beta_5 \)) are significant under all three regimes, i.e., \( \text{Regime 1} \), \( \text{Regime 2L} \) and \( \text{Regime 2S} \). Except for two coefficients—the advertising coefficient in \( \text{Regime 2L} \) and the distribution coefficient in \( \text{Regime 2S} \)—all of the remaining seven marketing mix coefficients have signs that are expected (i.e., negative price coefficients and positive advertising and distribution coefficients). Next, we discuss the two marketing mix coefficients with unexpected signs.

First, let us consider the advertising coefficient in \( \text{Regime 2L} \). Despite the negative sign, its magnitude is much smaller (−0.0002) compared to its counterparts in \( \text{Regime 1} \) (0.30) and \( \text{Regime 2S} \) (0.16). In other words, the substantive impact of the negative advertising coefficient may be minor compared to that of either of the two positive advertising coefficients. That said, a possible reason for the unexpected negative sign could be the high amount of advertising expenditure typically incurred by a luxury brand during the quarter of its launch, which is typically followed by an expenditure reduction of approximately 30% over the next 10 quarters. This strategic pattern of temporally decreasing advertising expenditure for a luxury brand, combined with the temporally increasing sales pattern for the brand in the first 10 quarters, might explain the (mildly) negative estimated correlation.

Next, consider the distribution coefficient in \( \text{Regime 2S} \). This coefficient is negatively signed (−1.34), in contrast to its counterparts in \( \text{Regime 1} \) (0.195) and \( \text{Regime 2L} \) (1.525). A possible reason for the unexpected negative sign could be the observation that the distribution levels of standard SUVs remained more or less unchanged for 8 quarters after the launch of luxury SUVs before increasing rapidly by approximately 4% (presumably stealing share from the luxury SUVs). Unfortunately, this spurt in distribution reach happened a little too late to “stem the tide”, given that the SUV category had already reached its maturity stage and sales had even started to decrease mildly. This strategic pattern of temporally increasing distribution reach for a standard SUV brand, which co-occurs with a temporally decreasing sales pattern for the brand in \( \text{Regime 2S} \), might explain negative estimated correlation.

An explicit investigation of the two unexpectedly signed marketing mix coefficients (discussed in the previous two paragraphs) would necessitate a statistical specification of the decision processes by firms with regard to the setting of marketing mix variables for their brands in addition to the joint estimation of such a statistical model with our proposed demand model. This analysis goes beyond the scope and purview of the present study and, thus, is left for future research.

9. As expected, the estimated brand-to-brand IPC coefficient (i.e., \( \beta_3 \)) is positive and significant (3.8854), which means that the increased adoption of a brand has a positive effect on its future adoption. The additional effect of cumulative sales from Quarter 19 onwards is found to be negative for both \( \text{Regime 2L} \) and \( \text{Regime 2S} \) and to yield a decreased, but still positive, impact of brand-to-brand IPC after the structural break. A possible reason could be the fact that SUVs were better understood by Quarter 19, with sales being less likely to be influenced by IPC compared to earlier periods.

10. The estimated order of entry coefficient is negative (−0.0707), which suggests that the early entrants have higher brand sales...
growth than later entrants (after accounting for the effects of lagged cumulative sales). This is consistent with the previous findings (Bowman & Gatignon, 1996).

11. Gasoline price is found to have a significant negative effect (−0.0018) on overall SUV demand. In other words, consumers shy away from buying SUVs when gasoline prices increase. This finding makes intuitive sense considering that SUVs are known to be gasoline guzzlers.

12. Our data analysis begins at the third quarter of 1990. As explained earlier, we chose this period as the starting point due to both the seeming onset of a new diffusion cycle in the SUV category (as evident in Fig. 2) and the trend for families with minivans to upscale to SUVs in the 1990s. We investigate the sensitivity of our parameter estimates to our choice of starting point in the following way: We now assume that category diffusion had already commenced prior to our starting point and, accordingly, let F0, the cumulative sales function evaluated at the starting point, take a positive value (instead of 0). Next, we re-compute the baseline hazard (first term on the RHS of Eq. (2)) and re-estimate the model. In Table 3, we report the estimates obtained using F0 = 0.1. Comparing Tables 2 and 3, we see that the value of the innovation coefficient p decreases considerably when F0 is changed to 0.1, which makes intuitive sense because initial adoptions are captured primarily by the parameter p. However, there are hardly any differences in the remaining coefficients. In this sense, our key findings seem to be immune to our choice of starting point for category diffusion.

13. Finally, we investigate whether our postulated structural break in model parameters in 1995 is warranted in the data. To do so, we re-estimate our proposed model over the full study period without allowing for a structural break during 1995 and also without distinguishing between luxury and standard SUVs. Table 4 reports the results based on this benchmark model. In terms of its ability to predict the market shares of the top four brands, the predicted RMSE, averaged across all 42 periods, is 0.037358 for the benchmark model and 0.035446 for the proposed model. In other words, the proposed model performs better than the benchmark model. More importantly, in terms of the estimated parameters, we notice two key changes in Table 4 compared to Table 2. The inclusive value coefficient (γ), which represents the collective impact of brand-level activities (that includes brand-level IPC) on category sales, more than doubles in value (from 0.3372 to 0.7328) and the brand-to-brand IPC coefficient (β4) drops to less than half its original value (from 3.8854 to 1.2753). In other words, if the structural break due to the introduction of luxury SUVs is ignored in the estimation, the brand-to-category IPC effects are overestimated while the brand-to-category effects are underestimated.

Next, we discuss some managerial take-home messages associated with our key findings.

5.4. Key managerial takeaways

Our empirical findings show the following patterns:

(a) In the SUV market we analyzed, a brand’s sales in quarter t were affected by brand-to-brand IPC generated by previous adopters of the brand up to quarter t − 1 (as captured by the parameter β4) and the brand-level marketing mix variables in quarter t (as captured by the parameters β3, β2 and β1) and

(b) The category sales in quarter t were affected by category-to-category IPC (as captured by the parameter q), brand-to-category IPC and marketing mix efforts (as captured together by the parameters β1 through β4 acting through γ). The latter two, in turn, represent the influence of prior adopters up to quarter t − 1.

Although it is difficult to generalize these SUV-findings to all the consumer durables, the fact that IPC plays such a strong role in general in the adoption of many consumer durables at the category level (as evidenced by the numerous research articles in the diffusion literature) gives credence to the possibility that the three types of IPC do exist and play an active role in the diffusion of the majority of consumer durables.

### Table 3

<table>
<thead>
<tr>
<th>Period</th>
<th>Regime 1</th>
<th>Regime 25 (std. SUV)</th>
<th>Regime 2L (Lux. SUV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 1990–Dec 1994</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
</tr>
<tr>
<td>Jan 1995–Dec 2000</td>
<td>0.0699</td>
<td>0.0741</td>
<td>0.0741</td>
</tr>
<tr>
<td>Category-level parameters</td>
<td>γ=0.000043</td>
<td>0.000043</td>
<td>0.000043</td>
</tr>
<tr>
<td>p=0.06945</td>
<td>0.0741</td>
<td>N.A.</td>
<td>1.2141</td>
</tr>
<tr>
<td>Gas price</td>
<td>−0.0045</td>
<td>−0.0045</td>
<td>−0.0045</td>
</tr>
<tr>
<td>Brand-level parameters</td>
<td>β1=2.0216</td>
<td>N.A.</td>
<td>1.2141</td>
</tr>
<tr>
<td>β2=0.8645</td>
<td>0.1213</td>
<td>N.A.</td>
<td>1.2141</td>
</tr>
<tr>
<td>Price</td>
<td>−0.8787</td>
<td>−0.7966</td>
<td>−0.7791</td>
</tr>
<tr>
<td>Advertising</td>
<td>0.3128</td>
<td>0.1679</td>
<td>−0.0043</td>
</tr>
<tr>
<td>Distribution</td>
<td>0.3108</td>
<td>−1.2633</td>
<td>1.2927</td>
</tr>
<tr>
<td>Lag Cum Sales</td>
<td>4.2133</td>
<td>4.2133 and 4.2133 and</td>
<td></td>
</tr>
<tr>
<td>Order of entry</td>
<td>−0.0690</td>
<td>−0.0690</td>
<td>−0.0690</td>
</tr>
<tr>
<td>Adv carryover</td>
<td>0.21</td>
<td>0.21</td>
<td>0.21</td>
</tr>
</tbody>
</table>

LL = −88,464.998; (Category incidence LL = −60,507,657, Brand choice LL = −27,957,342), BIC = 176,930,500.

### Table 4

<table>
<thead>
<tr>
<th>Period</th>
<th>Regime 1</th>
<th>Regime 25 (Jul. 1990–Dec 2000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category-level parameters</td>
<td>γ=0.7328</td>
<td>0.0020</td>
</tr>
<tr>
<td>p=0.0542</td>
<td>0.0016</td>
<td></td>
</tr>
<tr>
<td>Brand-level parameters</td>
<td>β1=1.0241</td>
<td>0.9790</td>
</tr>
<tr>
<td>β2=0.5880</td>
<td>0.8892</td>
<td></td>
</tr>
<tr>
<td>Price</td>
<td>0.5435</td>
<td>0.6577</td>
</tr>
<tr>
<td>Distribution</td>
<td>1.504</td>
<td>1.2753</td>
</tr>
<tr>
<td>Lag cum sales</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>Order of entry</td>
<td>0.6450</td>
<td>0.0577</td>
</tr>
<tr>
<td>Adv carryover</td>
<td>0.21</td>
<td></td>
</tr>
</tbody>
</table>

LL = −94,614.577; (Category incidence LL = −56,539,876, Brand choice LL = −28,054,701), BIC = 160,229,420.

All estimates are statistically significant at the 0.01 level. M was assumed to be 18,215,203. However, changing M by 20% did not significantly change the estimates. The advertising carry-over coefficient was identified through a grid search (based on the maximum likelihood criterion).
5.5. Simulation demonstrating the roles of IPCs

The roles of the three types of IPC effects and their corresponding parameters are illustrated in Fig. 1. As mentioned in Section 4, it is not possible to simply compare the estimates of the respective parameters and draw inferences about the relative impact of the three IPCs given the recursive and dynamic nature of the impact of IPC. We use a simulation to demonstrate how a practicing manager can use our proposed framework to develop a comprehensive understanding of the various roles performed by IPC.

We derive the sales implications of 10% price reduction of the Ford Explorer in Quarter 24 (where the chosen quarter and price reduction are both arbitrary). Consistent with the behavioral foundations of our proposed model, we explore the implications of the price reduction using two simulations run in sequence. First, we document the impact of the price reduction on brand sales that arises purely from the brand choice component of our proposed model (i.e., due to \( \beta_1, \beta_2, \beta_3, \beta_4 \)). Second, we document the additional impact of the price reduction on brand sales that arises from the category sales component of our proposed model (i.e., due to \( \gamma \)). Before we undertake these two simulations, we first simulate brand sales for all brands over 42 quarters using our estimated parameters (and keeping all explanatory variables at their observed values in the data). These “control” simulations are performed in two ways. First, we assume that \( \gamma = 0 \) (i.e., brand-level activities do not impact category sales). Second, we relax the restriction on \( \gamma \) (i.e., brand-level activities impact both brand share and category sales). The first manipulation yields Baseline Sales Growth: Brand Effect Only and the second yields Baseline Sales Growth: Brand Plus Category Effects, for all brands. For the following two “test” simulations, we assume that the leading brand, Ford Explorer, reduces its price in quarter 24 by 10% (assuming that all other variables for all brands remain unchanged). Again, as with the control simulations, depending on whether we assume \( \gamma = 0 \), two cases arise for the test simulations.

5.5.1. Simulation 1 (\( \gamma = 0 \)): Impact of brand-to-brand IPC (\( \beta_4 \))

Under Baseline Sales Growth: Brand Effect Only, the sales of Ford Explorer in Quarter 24 are estimated to be 72,025 units. A price reduction of 10% on Ford Explorer in Quarter 24 leads to a predicted increase in Ford Explorer sales during Quarter 24 of 10,330 units (i.e., 14.3%). Because category sales are not allowed to change under this simulation, this predicted sales increase for the Ford Explorer occurs solely at the expense of the current sales of its competing brands. These current effects depend on the parameter \( \beta_1 \).

Next, the increased sales for the Ford Explorer in Quarter 24 leads also to an increase in future sales (from Quarters 25 to 42) for the Ford Explorer due to the estimated positive brand-to-brand IPC effects, as captured by the parameter \( \beta_4 \). For example, the sales increase in Quarter 25 (57 units) arises because of the positive IPC produced by the 10.330 additional adopters of Ford Explorer in Quarter 24. In fact, the total predicted increase in the future (i.e., from Quarters 25 to 42) sales of the Ford Explorer is 1094 units. This represents the “brand-to-brand IPC” effect.

To summarize, a price reduction of 10% on the Ford Explorer in Quarter 24 results in (a) a current sales increase of 10,330 units in Quarter 24 (due to marketing mix effect, \( \beta_1 \)) and (b) a future sales increase of 1094 units in Quarters 25 to 42 (due to “brand-to-brand IPC” effects, \( \beta_4 \)). In percentage terms, the future sales increase is a sizeable percentage (10.6%) of the total sales increase of the promoted brand, which underscores the importance of IPC in this product category. Similarly significant dynamic spillover effects of price promotions have previously been obtained in packaged goods categories (see, for example, Seetharaman, 2004).

Re redoing this simulation by assuming that the price reduction on the Ford Explorer occurred in Quarter 7 (instead of Quarter 24) leads to the finding of the current sales of the Ford Explorer increasing by 5313 units (11.9% of the baseline sales) in Quarter 7 and the future sales of the Ford Explorer increasing by 2347 units in Quarters 8–42.

We redo this simulation for two other brands in the product category: Chevy Blazer, a standard SUV and Grand Cherokee, a luxury SUV. For each case, we uncover sizeable impacts of IPC effects.

We repeat this simulation exercise using multiple draws of the estimates from their joint sampling distribution (i.e., by using both the estimates and the estimates of their standard errors). Doing so yields a 99% confidence band for the predicted sales changes. These results are summarized in Table 5 (with the 99% confidence bands in square brackets). Simulations such as these can inform brand managers on potentially substantive IPC effects in terms of stimulating additional sales for their brands at the expense of their competitors.

5.5.2. Simulation 2: additional impact of \( \gamma \) on brand sales

It is useful to note that, under our proposed model, category sales growth is affected by both category-to-category IPC (captured by the parameter \( q \)) and brand-to-category IPC (captured by the parameter set \( \{\beta_4, \gamma\} \)). Under Simulation 1, we had ignored the latter IPC effect by setting \( \gamma = 0 \). Under the current simulation, however, we relax this restriction and let \( \gamma \) take its estimated value, i.e., 0.3372.

Under Baseline Sales Growth: Brand Plus Category Effects, the sales of the Ford Explorer in Quarter 24 are estimated to be 105,451 units. A price reduction of 10% on the Ford Explorer in Quarter 24 leads to a predicted increase in Ford Explorer sales of 16,941 units (i.e., 16.1%) in Quarter 24. The additional predicted sales increase of 6611 units in Quarter 24 compared to the sales increase documented in Simulation 1 (i.e., 16,941 units − 10330 units = 6611 units) arises due to the increase in category sales which, in turn, benefits the sales of the Ford Explorer. The 6611 units amount to a 64% increase over the sales in Simulation 1 (i.e., 6611/10,330). Therefore, modeling the category effect (using \( \gamma \)) in response to the price promotion of a brand is of substantive importance to a brand manager.

As in Simulation 1, the increased sales of the Ford Explorer in Quarter 24, in turn, increases the future sales of the Ford Explorer. The total predicted increase in future sales of the Ford Explorer (i.e., from Quarters 25 to 42) is 1305 units. By explicitly modeling the effects of brand-to-category IPC, we are able to conclude that a future sales increase of 211 units (i.e., 1305 units − 1094 units) will accrue to the Ford Explorer.

We redo this simulation by setting Quarter 7 as the period of price reduction for the Ford Explorer. We also redo this simulation for two other brands, Chevy Blazer and Grand Cherokee. For every case, we observe substantively sizeable brand-to-category IPC effects.

We repeat this simulation exercise using multiple draws of the estimates from their joint sampling distribution (i.e., by using both the estimates and the estimates of their standard errors). Doing so yields a 99% confidence band for the predicted sales changes. These results are summarized in Table 6 (with the 99% confidence bands presented in square brackets). In summary, Simulation 2 illustrates that ignoring the effects of brand-to-category IPC, as typically performed in previous research, can lead to serious underestimation of the total impact of price reduction on the sales of a brand. This is reminiscent of similar illustrations using static models of demand for packaged goods brands (see, for example, Chib, Seetharaman, & Strijnev, 2004).

6. Conclusions

IPC plays an important role in the market adoption of new products. However, the extant literature has focused largely on category-level IPC. KBK (2000) show that all the previous adopters

Our managerial simulations work under the assumptions that model parameters are invariant and that competitors do not react to the policy change (i.e., price promotion). Neither of these assumptions may be valid. However, testing their validity (or lack thereof) is beyond the scope of this research.
of a category affect the sales of a brand through IPC. In a recent and important paper, LMP (2009) extend this result to show that, in a duopoly market, the sales of a brand are affected by IPC from both its own previous adopters and previous adopters of the competing brand. In our study, we further generalize and extend the brand-level results of KBK (2000) and LMP (2009) using a brand choice model that is derived from utility maximization primitives. We show that there are three components of IPC affecting the sales of a brand. The first is the influence from the previous adopters of the brand, which we call brand-to-brand IPC. The second is the influence from the previous adopters of the brand that affects the category as a whole, which, in turn, affects the sales of a brand, which we call brand-to-category IPC. The third is the influence from the previous adopters of the category (i.e., regardless of the brand) that affects the category as a whole, which, in turn, affects the sales of a brand, which we call category-to-category IPC. The third component is typically the only type of IPC that has been modeled in the literature on diffusion models since the seminal work by Bass (1969). Using a careful model-based accounting of IPC effects, we disentangle the influences of the three IPC components while allowing for the effects of marketing variables. Our model incorporates repeat purchases while controlling for the heterogeneous order of entry across brands and gasoline price.

Ours is the first study to model three separate types of IPC effects within a brand-level diffusion framework. In doing so, we generalize category-level diffusion models such as the Bass Model (which handles category-level IPC) to additionally handle brand-level demand and brand-level IPC effects. Another advantage of the proposed framework is that it can handle multiple brands with heterogeneous entry times.

As a natural by-product of our modeling approach, we find that the adoption process of the typical consumer in the SUV product category seems to follow a two-stage process. In other words, category incidence is followed by brand choice. In behavioral terms, this means that the consumer first decides to buy the new product (“Should I buy an SUV?”) before deciding on the specific brand (“Which brand of SUV to buy, Ford Explorer or Jeep Grand Cherokee?”). However, given that only aggregate data were used in this study, further insight could be derived using individual-level data. We leave this investigation for future research.

As shown in our modeling framework, obtaining a nuanced understanding of a brand’s IPC (in terms of both brand-level and category-level components) and its separate influences on two consumer decision-making stages will provide measurable value to brand managers in the process of estimating the sales impact of their IPC campaigns. Such modeling expertise will acquire special significance in light of the increasing use of IPC marketing in the future.

Appendix A. Model derivation based on the utility-theoretic framework

Consider a consumer observed during quarter \( t \). During this quarter, suppose we observe a binary outcome variable \( y_t \) that takes the value 1 if the consumer buys an SUV during that quarter and 0 otherwise. This variable captures the customer’s category-level purchase decision ("incidence"). Suppose a category-level purchase occurs (i.e., \( y_t = 1 \)), let us say that we observe a multinomial outcome variable \( \gamma_t \) that takes the value \( j \) if SUV brand \( j \) is bought by the consumer. This variable captures the consumer’s brand-level purchase decision ("brand choice"). Our goal is to model the outcome variables \( \{y_t, \gamma_t\} \) on the basis of the prices \( P_{ij} \), advertising \( A_j \), and distribution strategies \( D_j \) followed by various available brands during quarter \( t \). We develop this joint model of incidence and brand choice sequentially below.

Category incidence model

Here, we develop a model of the binary outcome \( y_t \). Let \( z_t \) denote the (indirect) utility derived by a consumer from buying an SUV during quarter \( t \). This quantity is expressed as follows:

\[
    z_t = \ln \left( \frac{1 - F_{t-1}}{1 - F_t} \right) + \gamma C_{At} + \delta \cdot \text{GasPrice}_t + \epsilon_t, \quad (A1)
\]

where the second term on the right hand side captures the category attractiveness created by brand-level marketing activities and the IPC generated by prior adopters of each brand (i.e., Brand-to-Category IPC). The first term captures Category-to-Category IPC (where \( F_t \) stands for the distribution function that is associated with the consumer’s time of adoption for the SUV). GasPrice refers to the price of gasoline during

Table 5

<table>
<thead>
<tr>
<th>Sales increase due to price cut [impact of ( \gamma )]</th>
<th>Ford Explorer</th>
<th>Grand Cherokee</th>
<th>10% Price cut in Quarter 7 on</th>
<th>Ford Explorer</th>
<th>Chevy Blazer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales increase due to price cut [impact of ( \gamma )]</td>
<td>10,330 units in Quarter 24 [10,215, 10,441]</td>
<td>10,152 units in Quarter 24 [10,059, 10,245]</td>
<td>5313 units in Quarter 7 [5254, 5371]</td>
<td>2249 units in Quarter 7 [2224, 2274]</td>
<td></td>
</tr>
<tr>
<td>Sales increase due to increased brand-to-brand IPC [impact of ( \lambda )]</td>
<td>1094 units in Quarters 25–42 [1052, 1135]</td>
<td>1094 units in Quarters 25–42 [1058, 1131]</td>
<td>2347 units in Quarters 8–42 [2293, 2401]</td>
<td>433 units in Quarters 8–42 [424, 442]</td>
<td></td>
</tr>
</tbody>
</table>

\( \gamma = 0 \), i.e., category sales and growth are not affected by price cuts. Numbers in square brackets provide a 99% confidence band.

Table 6

<table>
<thead>
<tr>
<th>Sales increase due to price cut [impact of ( \gamma )]</th>
<th>Ford Explorer</th>
<th>Grand Cherokee</th>
<th>10% Price cut in Quarter 7 on</th>
<th>Ford Explorer</th>
<th>Chevy Blazer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales increase due to price cut [impact of ( \gamma )]</td>
<td>16,941 units in Quarter 24 [16,859, 17,023]</td>
<td>14,988 units in Quarter 24 [14,928, 15,047]</td>
<td>10,777 units in Quarter 7 [10,671, 10,883]</td>
<td>3929 units in Quarter 7 [3907, 3951]</td>
<td></td>
</tr>
<tr>
<td>Sales increase due to increased IPC [impact of ( \gamma )]</td>
<td>1305 units in Quarters 25–42 [1239, 1371]</td>
<td>1165 units in Quarters 25–42 [1111, 1218]</td>
<td>6218 units in Quarters 8–42 [6162, 6273]</td>
<td>874 units in Quarters 8–42 [869, 879]</td>
<td></td>
</tr>
<tr>
<td>% sales increase due to ( \gamma ) [&quot;brand-to-category IPC&quot;]</td>
<td>64.0% in Quarter 24</td>
<td>47.6% in Quarter 24</td>
<td>103.0% in Quarter 7</td>
<td>75.0% in Quarter 7</td>
<td></td>
</tr>
</tbody>
</table>

Numbers in square brackets provide a 99% confidence band.
quarter t and \( z_i \) is a stochastic component that is distributed iid (across t) Logistic with location 0 and scale 1.

The category-level outcome \( y_i \) is determined according to the sign of \( z_i \), as follows:

\[
y_i = I[z_i > 0],
\]

where \( I[A] \) is the indicator function taking the value 1 when the event A occurs and the value 0 otherwise. In other words, the consumer buys an SUV if the category utility exceeds his reservation utility (normalized to zero for identification purposes). This assumption yields the following category-level purchase probability (“incidence probability”) for the consumer.

\[
P_t = \Pr(y_t = 1) = \frac{e^{z_t}}{1 + e^{z_t}},
\]

Ignoring the second and third terms on the right hand side of Eq. (A1), Eq. (A3) reduces to the following:

\[
\Pr(y_t = 1) = \frac{\ln \left( \frac{1 - e^{z_t}}{e^{z_t}} \right)}{1 + \ln \left( \frac{1 - e^{z_t}}{e^{z_t}} \right)},
\]

which can be rewritten as

\[
\Pr(y_t = 1) = \frac{\ln \left( 1 + \frac{x - \frac{e^{z_t}}{1 + e^{z_t}}}{\frac{e^{z_t}}{1 + e^{z_t}}} \right)}{1 + \ln \left( 1 + \frac{x - \frac{e^{z_t}}{1 + e^{z_t}}}{\frac{e^{z_t}}{1 + e^{z_t}}} \right)}.
\]

Let \( x = \frac{e^{z_t} - \frac{e^{z_t}}{1 + e^{z_t}}}{1 + \frac{e^{z_t}}{1 + e^{z_t}}} \). Noting that x is positive and less than 1, assuming data of sufficiently small intervals of time, we apply the expansion of the natural log function in Eq. (A5) to obtain the following:

\[
\Pr(y_t = 1) = \frac{x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots}{1 + \frac{x}{2} + \frac{x^2}{4} + \frac{x^3}{6} - \cdots}.
\]

Note that x is a positive fraction and approaches a very small value as the time interval in the data set gets smaller and smaller. For sufficiently small intervals of time, the higher-order terms in both the numerator and denominator of Eq. (A6) become insignificant, thus yielding

\[
\Pr(y_t = 1) = \frac{x}{1 + x} = \frac{F_t - F_{t-1}}{1 - F_{t-1}},
\]

which is the discrete-time hazard that is associated with a continuous-time distribution function \( F_t \) (Seetharaman & Chintagunta, 2003).

We can assume any function for \( F_t \), but because the Bass Model (1969) has been proven as a robust model of category-level diffusion, we use the function \( F_t \) based on the Bass Model, which is as follows:

\[
F_t = \frac{1 - e^{-\left(p+q\right)t}}{1 + \frac{p}{q} e^{-\left(p+q\right)t}},
\]

where \( p \) and \( q \) stand for the innovation and imitation parameters, respectively. Doing so reduces Eq. (A4) to the discrete hazard to be associated with the BM.

**Category attractiveness effect**

The category attractiveness covariate, \( CAt_\text{t} \), which is contained within the first term on the right hand side of Eq. (A1), is an inclusive value measure that stands for the overall attractiveness of the SUV category—the composite measure of the marketing activities associated with all the available SUV brands—during quarter t. In other words, it is the function of the consumer’s indirect utilities for all available SUV brands during quarter t (as will be explained later), as shown below:

\[
CAt_\text{t} = \ln \left( \sum_{j=1}^{J_t} e^{V_{jt}} \right),
\]

where \( V_{jt} \) stands for the deterministic component of the consumer’s indirect utility for brand j during quarter t (this variable will be formally defined later) and \( j \) is the number of available SUV brands during quarter t. Our use of the inclusive value measure to represent \( CAt_\text{t} \) follows the same spirit as the nested logit model (Ben-Akiva & Lerman, 1985). The coefficient \( \gamma \), which represents the importance of the category attractiveness covariate in terms of influencing the consumer’s incidence decision in the SUV category, is expected to be positive.

This is our proposed incidence model. To the best of our knowledge, ours is the first attempt to integrate word-of-mouth effects, as represented in the BM, within a utility-theoretic framework that also explicitly captures the effects of marketing variables.

**Brand choice model**

Here, we develop a model of the multinomial outcome \( y^*_t \). Let \( U_{jt} \) denote the (indirect) utility derived by the consumer from buying SUV brand j during quarter t. We assume that this utility can be expressed as a function of the entire set of brand-specific covariates facing the consumer, in the following way.

\[
U_{jt} = \alpha_j + \beta_1 \cdot Price_{jt} + \beta_2 \cdot AdStock_{jt} + \beta_3 \cdot Dist_{jt} + \beta_4 \cdot CumSales_{jt-1} + \beta_5 \cdot OE_j + \epsilon_{jt},
\]

where the deterministic component of the indirect utility is given as follows:

\[
V_{jt} = \alpha_j + \beta_1 \cdot Price_{jt} + \beta_2 \cdot AdStock_{jt} + \beta_3 \cdot Dist_{jt} + \beta_4 \cdot CumSales_{jt-1} + \beta_5 \cdot OE_j,
\]

where \( Price_{jt} \), \( AdStock_{jt} \), and \( Dist_{jt} \) stand for the price, advertising-stock and distribution associated with brand j at time t, respectively, \( CumSales_{jt-1} \) stands for the cumulative sales associated with brand j up to \( t-1 \) and \( OE_j \) stands for the order of entry position for brand j. We assume that the errors \( \epsilon_{jt} = (\epsilon_{1jt}, ..., \epsilon_{njt}) \), are iid (across j and t) Gumbel distributed with location 0 and scale 1.

The brand-level outcome \( y^*_t \) is now determined in the usual way by the principle of maximum utility. We observe the outcome \( y^*_t = 1 \) when the utility of the jth SUV brand exceeds that of the remaining SUV brands. Specifically, we have the following:

\[
y^*_t = j \iff u_{jt} > \max_{k \neq j} u_{kt}.
\]

One would expect that, all else being equal, the effect of higher prices will be to depress utility (and, hence, to reduce the consumer’s probability of purchase for the SUV brand), whereas the effect of higher advertising (or distribution) and cumulative brand sales will be to increase utility. This yields the following brand choice probability at the consumer level:

\[
P_{jt} = \Pr(y^*_t = j) = \frac{e^{V_{jt}}}{\sum_{k=1}^{J} e^{V_{kt}}},
\]
which is the familiar MNL model (McFadden, 1974). It is useful to note that the summation in the denominator of the above equation runs over the number of available SUV brands during quarter \( t \), which is not necessarily a constant quantity over time (as is usually the case with mature packaged goods categories), given that new brands enter the category at various times and existing brands exit from the market at various times.

### Appendix B. Repeat purchases

The model and the associated estimation procedure, as developed thus far, pertain to first-time purchases only. However, because our study period covers a period of 10 years, it is possible for observed sales in later years to include both first-time purchases and replacement purchases. To handle this issue, we impute the number of replacement purchases in each quarter then subtract these imputed replacement purchases from the observed sales in each quarter to obtain a time series of first-time purchases. We use this time series of first-time purchases to fit the proposed model. To estimate model parameters, we adopt the following iterative procedure.

#### Step 0: Set first-time purchases = observed sales.

#### Step 1: Using first-time purchases, we fit our proposed model of incidence and brand choice outcomes to the data. This yields an estimated parameter vector \( \hat{\beta} \).

#### Step 2: Assumption that replacement purchases start in the kth quarter (where \( k \) is exogenously pre-specified) and that the replacement process follows a diffusion process that is identical to the diffusion process governing first-time purchases (as estimated in Step 1), we generate replacement purchases.

#### Step 3: We deduct the replacement purchases generated in Step 2 from the observed sales to generate first-time purchases for the category, then use market shares of the brands to generate brand-level replacement purchases.

#### Step 4: Go to Step 1.

We cycle through the above steps until there is no change in the estimated parameter vector \( \hat{\beta} \) obtained in Step 1. In the estimation, we achieve convergence within 10 iterations.

A couple of caveats are in order here. First, our assumption that the same diffusion process governs first-time purchases and replacement purchases may be restrictive. It is possible that replacement purchases are governed by a lower value of \( q \) (because replacement buyers depend less than first-time buyers on word of mouth) and are more sensitive to the marketing mix (because the pressure to buy a second SUV is not as severe as the pressure to buy the first SUV). For the sake of model parsimony, we ignore these issues. In practice, however, depending on the available degrees of freedom in the data or the availability of survey information from existing (first-time and replacement) buyers, managers can modify our model to obtain a more flexible representation of replacement purchases. Second, the value of \( k \) is subjectively chosen. We assume that \( k = 17 \), i.e., replacement purchases cannot happen within 4 years of buying the first SUV and are governed by a probabilistic diffusion process thereafter. This assumption appears to be reasonable. In addition, the assumption resulted in a predicted share of 25% for repeat purchases in the observed sales at the end of our data, which also appears reasonable. We also tried alternative values of \( k \) (varying from 10 to 25) and the results turned out to be qualitatively similar to those reported here. Additionally, when we aggregated the generated repeat purchases to the predicted first-time purchases from the model, we obtained total category sales that fit the observed sales data very well.

### Appendix C. Estimation method

It is useful to note here that the inclusive value coefficient, \( \gamma \), captures the dependence between the consumer’s incidence and brand choice outcomes (see Eq. (1)). From the above exposition, it would seem that there are 4 parameters—\( \gamma, \delta, p, q \)—associated with the incidence model and \( J + 5 \) parameters—\( \{ \alpha_j ; j = 1, \ldots, J \}, \{ \beta_0, \beta_1, \beta_2, \beta_3, \beta_4 \} \)—associated with the brand choice model (where \( J \) is the total number of SUV brands represented over our study period). However, that assumption is not quite correct. We will explain the dimensionality of our estimation problem here.

Because we have two regimes in our data (as explained in the empirical results section), we estimate a different set of brand choice parameters for Regime 2 compared to Regime 1 (also allowing one incidence parameter, \( q \), to change between regimes). Further, within Regime 2, we let brands (such as the Chevy Blazer) that stay in the same quality tier as in Regime 1 to have a different set of parameters compared to brands (such as Ford Explorer) that move upwards to the luxury tier. Doing so results in 3 sets of brand choice parameters, depending on whether the brand in question is in (1) Regime 1, (2) Regime 2—Standard, or (3) Regime 2—Luxury. This scenario means that the 3 parameters associated with the marketing variables \( -\gamma, \beta_0 \), and \( \beta_4 \) will be different under these 3 scenarios (leading to 12 parameters instead of 4). Lastly, we allow the 4 largest brands—Ford Explorer, Chevy Blazer, Jeep Cherokee, Jeep Grand Cherokee—to have brand intercepts (without intercept for Grand-Cherokee considering that it was not there in Regime 1) and set the intercepts for the remaining (smaller) brands to zero. Doing so brings the number of estimable parameters in the brand choice model to 20 (i.e., 7 brand intercepts, plus 12 slope parameters, plus 1 order-of-entry parameter). Taken together with the 5 parameters associated with the incidence model, the total number of estimable parameters in our proposed model is 25.

The sample likelihood function can be written as follows:

\[
L = \prod_{t=1}^{42} \left( 1 - P_{r,t} \right) M - \sum_{j=1}^{J-1} \left( \sum_{j=1}^{J} S_{j,t} \left( P_{j,t} \right) \right) - \sum_{j=1}^{J} S_{j,t} \left( \prod_{j=1}^{J-1} \left( I_{j,t} P_{j,t} \right) \right),
\]

where \( J \) is the total number of SUV brands represented in our data (over the study period), \( S_{j,t} \) is the observed sales for brand \( j \) during quarter \( t \), \( M \) is the market potential for SUVs, \( I_{j,t} \) is an indicator variable that takes the value 1 if brand \( j \) is available in the SUV market during quarter \( t \) and 0 otherwise. This likelihood function is maximized using gradient-based routines in Gauss.

### References


Parenthood.com (2012). http://www.parenthood.com/article-topics/minivan_or_suv_which_is_right_for_your_family.html


WSJ (May 20). online article: http://online.wsj.com/article/SB100014244052704448304575196172206855634.html