Accurate new product market performance forecasts are important for proper resource allocation decisions before market launch. Traditional forecasting methods are based on concept tests, conjoint analysis, or diffusion models. Pre-launch forecasts are plagued with rather limited accuracy rates. Internet prediction markets have provided promising results but can be applied only when participants know the products. All these methods build on rather strong assumptions concerning consumer information processes and decision-making. The observation of prelaunch new product discussions in Internet discussion forums may provide the forecast required for resource allocation decisions concerning the early introduction stage of new products. This study analyzes the forecasting potential of word-of-mouth data acquired from prelaunch new ski discussions. Data was collected from online-conversations in three discussion forums concerning ten different new ski models before market launch. The communication among forum members was passively observed for several months. Results show high coherence between pre-launch new product evaluations and product performance after market launch.

1. Introduction

Continuous product innovation is important for sustainable market success. New product launches may require substantial management attention as well as production, marketing and financial resources. Failing introductions of new products represent substantial risks for profitability. Allocating the right amount of resources to new products is of major importance. Having allocated too much of available resources to a new product that turns out to be less successful as expected diminishes profitability. Not being able to supply a surprisingly successful new product according to demand due to lacking resources reduces potential returns. Most often resource allocation decisions involve considerable lead-time to implement. Forecasting the market performance of new products after launch may be too late.

Despite the importance of timely and accurately forecasting the market performance of new products before launch, traditional new product forecasting methods often deliver inaccurate results. In a survey across different types of new products Kahn (2002) found an average forecasting accuracy of only 58 %. In light of the risk of losing opportunities when wrongly forecasting new product failure as well as misdirecting substantial resources in the case of inaccurately forecasting new product success, the importance of developing more precise forecasting tools appears evident.

Acknowledgements: The authors would like to thank two anonymous reviewers for their precise and constructive comments to the first version of the manuscript.
Traditional literature on new consumer product market performance forecasting suggests methods to elicit target customer self-assessment of purchasing intentions (Jagpal/Jedidi/Jamil 2007), conjoint analysis models for simulating future market shares (Wittink/Cattin 1989) and more or less complex statistical diffusion models (Peres/Muller/Mahajan 2010). Purchase intention measures suffer from their low correlation with subsequent purchasing behaviour (Morwitz/Schmittlein 1992). Conjoint analysis has proven to produce quite accurate forecasts of new product market shares in cases where consumers tend to extensive purchase decision-making processes and the product can be easily decomposed into a limited number of attributes (Karniouchina 2011). Diffusion models are typically based on aggregated sales data from the introduction phase of a new product and select a specific diffusion function based on theoretical assumptions about the interactions among consumers (Van den Bulte/Stremersch 2004). For consumer products with short time spans between introduction and takeoff to be timely, diffusion forecasts need to rely on parameters from previous product generations. These parameters may not allow adequately forecasting the diffusion of a new product because diffusion functions take different forms depending on social and competitive market dynamics (Golder/Tellis 2004).

Literature suggests that prediction markets are able to overcome these problems. Prediction markets aggregate relevant information from a multitude of diverse participants. Participants buy and sell stocks that represent a new product before market launch depending on their expectations concerning the future success of the product. The price of the stock represents the aggregated knowledge of all participants. Research has shown that the forecasts of new product success based on prediction markets outperform forecasts based on expert judgments, concept tests or conjoint analyses (Skiera/Spann 2004). However, the accuracy of prediction market results depends on a number of assumptions holding true (Ho/Chen 2007) which is not always the case (Karniouchina 2011).

As research on word-of-mouth in online discussion forums, blogs and other social media (Kozinets et al. 2010) has shown, consumers increasingly base their product choices on product evaluations posted by people who they consider being highly knowledgeable and more reliable than company promotion (Godes et al. 2005). Such opinion leaders most often are industry insiders, product category freaks or individuals highly experienced in a specific activity (Becheur/Gollety 2007). Companies provide them with information or new products long before general market launch. The opinion leaders post new product information or reports of their experiences in product or activity-related discussion forums on the Internet. The reports lead to interested exchanges among active consumers and tend to influence the purchase decisions of a substantial number of “tourists” (Kozinets 1999). Observing such processes allow marketers to gather information on the evolution of the dispersion, the volume and valence of ongoing discussions among potential customers before product launch without the need to rely on the behavioural assumptions of prediction markets. Research has shown that dispersion (Godes/Mayzlin 2004), volume (Liu 2006) and valence (Chevalier/ Mayzlin 2006) of Internet postings concerning movie pre-views have a significant effect on ticket sales. The question arises if such discussions in specialized online-forums can be used as a reliable predictor for the success or failure of other consumer products in the very early stage of market diffusion. So far no published research has attempted to answer this question.

The research described in the following is a first attempt to close that gap. The research contributes to the literature in suggesting and testing an alternative forecasting approach to be applied before the launch of new products. Forecasts build on data gathered by unobtrusive observation of real interactions among potential customers instead of surveys, assumptions about adoption behaviour or price mechanisms in simulated markets. Volume and valence of discussions in product-specific online forums concerning new consumer products are measured before launch for the purpose of forecasting the products’ market performance just after launch. The resulting forecasts appear to be highly accurate. The forecasts are coherent with first year sales figures and success evaluations of top managers responsible for sales.

The article starts with a literature review that discusses the potential of new product market performance forecasting approaches dominating the literature to accurately forecast new product performance in the stage of introduction. A discussion of new methodological opportunities provided through the development of Web 2.0 prepares the ground for the following presentation of an empirical study that observed the pre-launch communication in three online skiing discussion forums concerning ten new ski models. Volume of total views, threads (= publicly accessible discussions concerning a specific issue) and replies as well as the dominating valence of published opinions served to forecast market success or failure. Results show the usefulness of observing online product discussions before market launch for forecasting short-term new product success. The discussion points out the contribution of the research to the literature as well as managerial implications. A summary of limitations of study design and available data lead to suggestions for future research.

2. Literature Review

The literature suggests various ways of resolving the problem of forecasting the market performance for new products before launch. A still widely used approach is concept testing (Crawford/Di Benedetto 2006). In concept test surveys potential customers indicate their level of purchase intention concerning the concepts of one or
more products presented to them (Ulrich/Eppinger 2004). Concept testing is a rather simple method to forecast product demand. The method can be applied where simulated or full-scale test markets for predicting new product demand are not feasible. However, the application of traditional concept testing to forecasting new product sales suffers from several weaknesses. Most published studies used scalar self-stated intentions data to predict product demand. Scalar measures do neither allow estimating measurement error nor directly translating intentions into choice probabilities (Jagpal/Jedidi/Jamil 2007). Interviews are prone to demand artefacts as well as situational influences. Additionally, consumers tend to follow the example of early adopters (Van den Bulte, C./Joshi, Y.V. 2007). In product test situations, such imitators are unable to reliably indicate their purchasing behaviour for products that have not been launched so far. Potential cannibalization effects cannot be assessed. Extant concept-testing methods neither consider the impact of word-of-mouth, nor the effect of marketing activities on product diffusion. Consequently, most resulting forecasts exhibit substantial error (Peng/Finn 2010).

Conjoint analysis has been extensively used to predict market shares of new products (Wittink/Cattin 1989). Conjoint analysis assumes that consumers base their preferences on the assessment of product attribute utilities and related tradeoffs. Conjoint designs present combinations of product attributes to consumers and ask the respondents for an evaluation of attractiveness or preference. The estimate of potential market share results from the comparison of the utility of the new bundle of product features to the utilities of feature combinations existing on the market. Therefore, conjoint analysis is appropriately used for new product market performance forecasts when consumers are ready and able to go through extensive decision-making processes (Mühlbacher/Botschen 1988) and when the number of product attributes is limited. Product categories where habitual or spontaneous buying behaviour predominates or where usage experience, emotions, imagination or symbolic value lead to a more gestalt-like evaluation of products are less well suited for new product performance forecasting with conjoint analysis (Karniouchina 2011).

Diffusion models strive to predict the number of adopters of a new product as an explicit function of time or space. The Bass diffusion model (Mahajan/Muller/Bass 1990), for example, provides information on the development and speed of new product sales in a social system over time. Bass (1969) uses aggregated sales data to infer the model’s coefficient of internal influence, which expresses the assumed level of word-of-mouth. Innovators first adopt new products. Followers learn about the new product through word-of-mouth and imitate earlier adopters. As research on new product takeoff has shown, takeoff can result from consumer word-of-mouth or imitation but is also influenced by consumer heterogeneity in terms of price sensitivity and risk avoidance. In most cases a mix of these factors will influence the success of a new product in the early stage from market launch to takeoff. The resulting diffusion curves can depart significantly from the typical bell-shaped sales curve of the Bass model (Golder/Tellis 1998; Van den Bulte/Joshi 2007).

Over the years, the Bass diffusion model has been refined and extended (Chandrasekarar/Tellis 2007). To overcome the model’s overemphasis on word-of-mouth communication researchers have introduced other types of social interactions, such as observed social signals (Van den Bulte/Joshi 2007) or network externalities (Stremersch/Binken 2009). Agent-based modelling serves to relax the implicit assumption that the population of consumers is homogeneous and consumers are fully connected (Delre et al. 2010; Goldenberg/Libai/Muller 2010). These more sophisticated diffusion models provide a better fit to empirical diffusion data. However, as Peres/Muller/Mahajan (2010) point out, the interface between the individual level and the aggregate level in agent-based models lacks a closed formulation and needs further exploration. Decker/Gnibba-Yukawa (2010) stress the need of agent-based models to estimate additional parameters. Thus, agent-based models are less suited for sales forecasts based on limited time series data from new product “incubation time” (Kohli/Lehmann/Pae 1999).

Golder/Tellis (1997), Agarwal/Bayus (2002) and Tellis/Stremersch/Yin (2003) suggested models that are able to forecast sales takeoff based on data from the introduction stage. These models seem suitable for consumer durable innovations, which typically take substantial time before taking off. For simply modified consumer products with rather short time spans before takeoff, forecasts must be made before launch for enabling managers to provide the needed resources. Forecasts using diffusion models before product launch have to rely on estimates from analogous products. Researchers of entertainment products have successfully developed industry specific models to forecast market performance before product launch (Eliahsberg et al. 2000; Lee/Boatwright/Kamakura 2003). However, as Chandrasekarar/Tellis (2007) point out these models may not be suitable for other consumer products.

With the advent of Web 2.0, researchers are able to directly assess consumer expectations concerning a new product’s potential market success through prediction markets (Skiera/Spann 2004; Spann et al. 2009). In contrast to the measurement of individual preferences through concept tests or conjoint analysis, prediction markets rely on the wisdom of consumer crowds (Koznets/Hemetsberger/Schau 2008). In a game-like environment participants engage in stock trading. The stock represents a new product that is to be launched. The price traders are willing to pay for a stock represents their expectation concerning the market success of a new product in view of other traders’ expectations which are expressed by the current market price. Traders may learn
from the behaviour of fellow traders and can improve their forecasts step by step. When the product is launched, traders are compensated according to the value of their predictions (Dahan/Hauser 2002). The more accurate the forecast the higher the rewards participants receive. Empirical research has shown that the forecasts of new product success based on prediction markets outperform forecasts based on expert judgments, concept tests or conjoint analyses (Skiera/Spaß 2004).

The superior forecasting performance of prediction markets depends on the fulfilment of several conditions (Ho/Chen 2007): (1) those who have more accurate information than others are more confident and buy more stock. (2) Traders have different mental models, which are precisely expressed by the price metric. (3) Participants have independent sources of knowledge. (4) Uninformed participants learn from the better informed, and (5) pooling of information across participants leads to increasingly accurate information. As research has shown the amount of stock bought is not always an unbiased indicator for the availability of accurate information. Dahan et al. (2011) found that traders tend to exhibit a self-preference bias when trading. Where the performance of traders cannot be measured based on actual observable results, some people tend to gamble hoping to make excessive returns. Karmiouchina (2011) pointed out that highly visible conspicuous information may be over-utilized whereas information that is difficult to obtain and to process may not be fully reflected in stock prices. This bias may be particularly pronounced if the experience of the participants is not effectively relevant for the new product at hand. The pooling of opinions provides reliable results only if there are a sufficiently large number of traders who do not base their judgment on the same (biased) sources of information. In this respect, Delre et al. (2010) found that the persuasive power of VIPs on other consumers is not stronger than the influence of peers. However, VIPs have a greater capacity of informing other consumers. The literature on herding behaviour (Salganik/Dodds/Watts 2006) has shown, that pooling of individual information may lead to more homogeneous instead of better information.

With the spread of social media product diffusion processes increasingly take place in social networks on the web. Researchers have found that online product reviews have substantial influence on consumer behaviour (Chevalier/Mayzlin 2006; Senecal/Nantel 2004) and statistically significant relationships with future sales (Belvaux/Flories 2009; Chintagunta/Gopinath/Venkataraman 2010; Duan/Gu/Whinston 2005; Godes/Mayzlin 2004; Liu 2006). Dellarocas/Awad/Zhang (2006) showed that the early volume of online reviews on Yahoo! Movies (that is the amount of word-of-mouth disseminated) can be used as a proxy of sales, allowing revenue forecasting to take place before early box office results are published. Liu (2006) found a direct relationship of the volume of word-of-mouth concerning a new motion picture with box office results. Forecasts based on the volume of online ratings of music albums allowed generating reliable forecasts sooner than before (Dhar/Chang 2009). Researchers report mixed results concerning the impact of the valence of online word-of-mouth (Chen et al. 2004; Chevalier/Mayzlin 2006). Dhar/Chang (2009) found the valence of word-of-mouth in social media to have no impact on music sales. Duan/Gu/Whinston (2008) reported that the valence of word-of-mouth has an indirect influence on box office revenues, whereas Chintagunta/Gopinath/Venkataraman (2010) found the valence of posted opinions to have more influence on box office sales than the volume of posts. These findings are interesting because they relate to experience products with potentially very fast takeoff that do not allow performance forecasts with product concept tests, conjoint analysis or diffusion models. Major limitations of the available studies are their lack of generalizability because they focus on low risk experience products and data mainly gathered from review sites. Most studies used product assessments made after or very shortly before the launch of the new products (Chakravarty/Liu/Mazumdar 2010).

In social media networks new consumer product diffusion processes may start well before market launch. Companies exhibit their new products at trade shows months before market launch. Members of discussion forums, who work in the industry post verbal and visual information about these products. Once this information is introduced into the discussion forum, members start discussing the pros and cons of the changes to the products, make evaluations, and exchange opinions (Füller/Jawecki/Mühlbacher 2007). Companies provide their new products to experienced consumers, such as professional users, for testing purposes, or let shop owners try the new product during demo days, months before offering the product to the general public. Such users report their new product experience in details to interested discussion forum members who intensively discuss what they get told. Members asking the community for personal advice also initiate new product discussions (Sawhney/Verona/Prandelli 2005). Many less active visitors of discussion forums are interested “tourists” who base their purchase decisions on what they observe (Kozinets 1999).

Online discussion forums can be seen as live test markets to gain real-time consumer feedback on product innovations by unobtrusively observing the ongoing discussion processes (Füller et al. 2006). Unlike in consumer concept test surveys, data collection does not socially bias the findings. Consumers are not forced into using a number of product attributes when assessing a new product. There is no need for making assumptions about the intensity of word-of-mouth, the influence of signals or externality effects such as in diffusion models. In contrast to prediction markets, people need no monetary incentives to participate. They are interested in the product, the related activity or the other people participating in the forum.

https://doi.org/10.15358/0344-1369-2011-3-221

Das Erstellen und Weitergeben von Kopien dieses PDFs ist nicht zulässig.
stead of simulating consumer information exchange and processing through stock and market price mechanisms based on assumptions about trader information use and decision making researchers can directly determine the role and influence of each individual actor in the discussion process and observe to what extent the forum members reach consensus in the evaluation of a new product.

In light of these advantages, research needs to be conducted for other product categories than motion pictures, books and music albums and other types of online gatherings than review sites in order to increase our understanding of the value of evaluative word-of-mouth on the Internet for forecasting new product market performance before launch. Therefore, the empirical study described in the following focuses on two questions:

1. Can the observation of word-of-mouth in online discussion forums well before market launch be effectively used for forecasting new product market performance in the introduction phase of the product?
2. What is the impact of volume and valence of pre-launch new product discussions in online forums on the market performance of new products?

3. Empirical study

3.1. Research Field

The observation and analysis of new product discussions in online forums for forecasting the success or failure of a new product before launch may be especially useful for products that are (1) difficult to assess, (2) cannot be evaluated before consumption or use and (3) have an impact on the social status ascribed by relevant others. Products that are rather complex or technical are difficult to assess before consumption or use, (4) market launch date, and (5) category of ski. The research exposed in the following focuses on alpine skis that is on experience products used in public. There exist a number of large, active online discussion forums interested in skiing activities and related experiences.

The forums differ in their structure and member bases. This diversity allows increasing the generalizability of findings by observing various discussion forums to cover a broad range of user characteristics. Members engage in active discussions of new products well before, during, and after market launch. New product discussions are archived and accessible. Finally, one of the authors disposes over considerable knowledge about and personal experience in skiing in general, the market, the people and the specific language (of insiders), which is valuable for data analysis and interpretation (Kozinets 2002).

3.2. Methodology

Netnography appears to be the most suitable method to investigate product-related discussions in online discussion forums. Netnography is a qualitative research methodology that adapts ethnographic research techniques to study online gatherings (Kozinets 2002). Whereas traditional ethnography focuses on the person as the unit of analysis, netnography observes communicative acts. The netnographic procedure followed in this study included four steps as suggested by Kozinets (1998; 2002) and described in the following.

3.2.1. Identification and selection of discussion forums

Step 1 first focused on the identification of online skiing discussion forums suitable for the research purpose. The search engine Google allowed finding a number of discussion forums. Search terms were “online skiing communities” and “ski forums”. Using Kozinets’ (2002) five general criteria for online community selection, the number of potential discussion forums could be narrowed down to three. The researchers aimed to find discussion forums with a high number of distinct threads as well as postings and views related to new ski models, detailed and descriptively rich data and substantial between-member interactions. Tab. 1 provides an overview of the characteristics of each selected discussion forum.

3.2.2. Selection of new ski models

Step 2 involved the familiarization with characteristics of the selected discussion forums. For eight weeks the researchers observed the communication among discussion forum members, analyzed group membership, products of interest and usage behaviour of forum members. The specific language used in each of the forums also needed some time for familiarization. Parallel to this familiarization ten different new ski models were selected for analysis.

Ski model selection followed five criteria: (1) degree of innovativeness, (2) level of discussion surrounding the ski, (3) existence of discussions before and after market launch, (4) market launch date, and (5) category of ski. The selection included products, which generated ample postings (thus creating a large word-of-mouth effect), as well as products, which were not discussed intensively.
The ten selected new ski models each belonged to one or more of the following product categories: (1) All-mountain, (2) Freestyle or (3) Freeride. Although some skis are designed specifically for one product category, it is quite common for ski manufacturers to build models for a broader range of use. Therefore, some skis belong to two different product categories, as can be seen in Tab. 2.

The selected new products represented varying degrees of innovativeness, and were introduced on the market in the USA in the winter season of 2005–06 (with the exception of the Atomic M:EX which was introduced one year earlier). Since the majority of community members were from the USA, only ski models, which had been launched in the USA, were available for selection. Since almost all ski models are sold on the North American market, this caveat did not present any limitations.

### Table 1: Profiles of selected discussion forums

<table>
<thead>
<tr>
<th>Forum</th>
<th>Registered Members</th>
<th>Main Focus</th>
<th>Discussion Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPICSki</td>
<td>&gt; 12.000</td>
<td>&quot;Expert community&quot;</td>
<td>&quot;Ski Gear Discussion&quot; with over 10.400 threads and 102.300 posts, one of the most popular forums.</td>
</tr>
<tr>
<td>NEWSchoolers</td>
<td>&gt; 87.300</td>
<td>Low member age, mostly teenagers</td>
<td>&quot;Ski Equipment&quot; with over 278.400 posts.</td>
</tr>
<tr>
<td>Teton Gravity Research</td>
<td>&gt; 3.360</td>
<td>Main focus is skiing but also some non-skiing related topics are discussed.</td>
<td>&quot;Tech Talk&quot; with over 5.000 threads and 62.000 posts.</td>
</tr>
</tbody>
</table>

### Table 2: List of analyzed new ski models

<table>
<thead>
<tr>
<th>No.</th>
<th>Brand</th>
<th>Ski model</th>
<th>Product category</th>
<th>Market launch (USA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Atomic</td>
<td>M:EX</td>
<td>All-mountain/Freeride</td>
<td>2004-05</td>
</tr>
<tr>
<td>2</td>
<td>Atomic</td>
<td>Sweet Daddy</td>
<td>All-mountain/Freeride</td>
<td>2005-06</td>
</tr>
<tr>
<td>3a</td>
<td>Atomic</td>
<td>Triplets HP</td>
<td>Freestyle</td>
<td>2005-06</td>
</tr>
<tr>
<td>3b</td>
<td>Atomic</td>
<td>Triplets Urban</td>
<td>Freestyle</td>
<td>2005-06</td>
</tr>
<tr>
<td>4</td>
<td>Armada</td>
<td>ANT</td>
<td>Freeride</td>
<td>2005-06</td>
</tr>
<tr>
<td>5</td>
<td>Line</td>
<td>Invader</td>
<td>Freestyle</td>
<td>2005-06</td>
</tr>
<tr>
<td>6</td>
<td>Nordica</td>
<td>Hot Rod Top Fuel</td>
<td>All-mountain</td>
<td>2005-06</td>
</tr>
<tr>
<td>7</td>
<td>Rossignol</td>
<td>Bandit B3</td>
<td>All-mountain/Freeride</td>
<td>2005-06</td>
</tr>
<tr>
<td>8</td>
<td>Salomon</td>
<td>1080 Foil</td>
<td>Freestyle</td>
<td>2005-06</td>
</tr>
<tr>
<td>9</td>
<td>Völkl</td>
<td>Unlimited AC4</td>
<td>All-mountain</td>
<td>2005-06</td>
</tr>
<tr>
<td>10</td>
<td>Völkl</td>
<td>Mantra</td>
<td>Freeride/All-mountain</td>
<td>2005-06</td>
</tr>
</tbody>
</table>

### 3.2.3. Data Collection

Step 3 consisted of forum observation and data collection. Data collection focused on conversations concerning the selected new ski models. All online discussions were read, and if centred on the selected new products, saved as a text file. Discussions concerning other topics than the selected new ski models, such as rather frequently occurring general brand related discussions in Newschoolers.com, were neglected.

To be able to analyze new product discussions over time, the researchers retrieved data in hindsight from the community.
Ski Manufacturer Data Sources for External Product Evaluation

<table>
<thead>
<tr>
<th>Ski Manufacturer</th>
<th>Data Sources for External Product Evaluation</th>
</tr>
</thead>
</table>
| Atomic           | • Personal interview with the International Marketing Manager, October 2005  
|                  | • Sales data for Atomic Triplets, M:EX, Sweet Daddy  
|                  | • Interview with the Atomic Team manager on Newschoolers.com |
| Armada           | • Telephone interview with the Sales Manager of Armada USA, January 2007, including sales data |
| LINE             | • Telephone interview regarding the market performance of Invader with the founder and CEO of LINE Skis, January 2007  
|                  | • Interview via email with “Jibij Pro Shop” retail store owner/manager, who caters to the freestyle market, November 2007  
|                  | • No sales data obtainable |
| Nordica          | • Telephone interview regarding the market performance of Hot Rod Top Fuel with the CEO of Nordica USA, January 2007  
|                  | • No sales data obtainable |
| Rossignol        | • No data obtainable |
| Salomon          | • Data regarding market performance of Foil 1080 via email from the Director Sales Operations and Service, February 2007 |
| Völkli           | • Interview via email with the Marketing Manager of Völkl Sport America & Tecnica USA, including sales data, September 2006 |

Table 3: Data sources for new ski model market performance

Data concerning the market performance of new ski models are not publicly available. They had to be gathered from the ski manufacturers. Generally, ski manufacturers do not share internal information such as sales data with people outside the company. However, one of the authors was able to access this sensitive data due to her involvement in some of the companies as a professional ski athlete. From three ski manufacturers, sales data of each new ski for the first winter season after market launch (the winter season 2005–06) within the North American market could be obtained. Collected sales data concern six out of the ten selected new ski models. In the remaining cases, the researchers collected expert information about the perceived success or failure of the ski in the marketplace directly from ski manufacturers’ top executives responsible for sales either through e-mail or through telephone interviews (Tab. 3). Executives were asked if actual sales in the first season had reached expectations, had surpassed expectations or had stayed below expectations. Interviewees had little incentive to provide false portrayals of their ski’s market performance, since the requested information concerned the previous winter season.

3.2.4. Data Analysis and Measures

Step 4 involved data analysis and interpretation. In total the researchers collected 938 threads, analyzed 542, and coded 398 threads. Almost 50% of all coded threads stemmed from Epicski, followed by Newschoolers.com with almost 38%, and TGR with a smaller contribution of just over 16%. These percentages reflect the amount and quality of new ski model-related information available in each discussion forum as well as the amount of discussions surrounding the selected ski models. A random sample of 22 threads revealed that the average length of threads in Newschoolers.com was just below 2,000 words, in TGR 2,647 words and in Epicski 3,245 words per thread. Analyzed data included product opinions posted by online forum members, photographs and researcher field notes. A difficulty in determining the number of total threads, replies, and views related to each new ski model came from the fact that in a large number of threads, discussants compared a variety of ski models. Therefore, one could not attribute the total number of replies and views in such a thread to one ski model alone. This would have made the ski seem more popular than it was. To avoid that potential bias the researchers decided to only include threads which where mainly devoted to the evaluation of one or more of the ten new ski models.

The qualitative data analysis software package NVivo (version 7) served to analyze the collected data. The researchers assigned the selected threads to individual folders called “Epicski”, “Newschoolers.com” and “TGR” in order to allow analyzing separately the threads found in each forum and to compare the threads across forums. Within each discussion forum, data was further organized into ski model sub folders.
Researchers considered a statement expressed by a forum member to be an “opinion” if it included a product evaluation. In total, 415 different online forum members contributed a total of 1351 consumer opinions concerning the selected ski models to the discussion. Researchers coded each consumer opinion according to the point in time (month/year) it was published and the author of the opinion, and independently sorted the opinions into four valence categories: (1) very positive, (2) positive, (3) neutral, and (4) negative. The “very positive” opinion category included product evaluations expressed by statements such as “I love this ski”, “this is the best ski I’ve ever been on”, or “I’ve never had so much fun on a pair of skis”. Discussants expressing “positive” opinions still considered the new ski model in a better than neutral manner. But their statements were less strongly emotive. For example, Billiam posted on Epicski: “The guy selling the Euro B2 on eBay claims that they have redesigned the tip to make it better in crud – and the width went from 76 to 78mm. Sounds like good changes to me. Makes it more versatile.” The “neutral” opinion category included statements in which forum members weighed advantageous product characteristics against disadvantageous product characteristics without making explicit judgments. Initially, researchers planned to use a fifth opinion valence category: “very negative”. This category was disregarded after discovering that most product opinions were positive, which confirms previous empirical word-of-mouth research (Chevalier/Mayzlin 2006).

Researchers split the data set into three overlapping parts, coded the overlaps first and compared the coding of these parts. Inter-coder reliability showed highly satisfactory levels. Where coders did not agree the categorization of opinions was discussed until reaching consensus. Based on the consensus of how to code, each researcher individually coded the rest of the attributed data.

To be able to assess the forecasting potential of pre-launch online forum discussions for new product market performance, the performance of new ski models had to be measured within each discussion forum before launch as well as in the market after launch. According to the literature (Duan/Gu/Whinston 2008), three indicators served for measuring new ski model performance inside each discussion forum: the dominant valence of the opinion (VO) forum members expressed when discussing a new ski model, the level of product interest of forum participants (PI), and the level of word-of-mouth (WOM).

The valence of opinion (VO) measure resulted from a comparison of the total number of positive versus negative opinions concerning a new ski model expressed by discussion forum members. According to the literature (Arndt 1967; Chevalier/Mayzlin 2006), negative opinions have a stronger effect on other consumers’ opinions than positive opinions. Additionally, discussion forums tend to show a bias in terms of more positive than negative statements (Godes et all. 2005). Taking these facts into consideration the researchers defined a new ski model as a success within a discussion forum when the amount of positive opinions exceeded 60 % of all opinions expressed. The fact that some new ski models earned a high level of negative opinions was taken into consideration by adding a second classification rule. If negative opinions concerning a new ski model exceeded 20 % of the total number of opinions expressed, positive opinions had to exceed the level of 70 % in order for the new ski model to be still classified as a success within the online discussion forum.

PI and WOM are distinct measures of the volume of word-of-mouth generated by a new ski model in a discussion forum. The PI level represented the total number of views of all threads in a forum largely dedicated to a particular ski. The more thread views a ski model received, the larger the product interest by the discussion forum, including “tourists”, who mostly read but do not post comments. The WOM level represented the total number of threads and replies within the threads which were generated within a discussion forum regarding a new ski model. The larger the total number of threads and replies the higher the level of WOM. Because WOM considered active contributions only this measure of word-of-mouth volume was more conservative than PI.

The researchers decided to calculate PI and WOM levels separately for each new ski model within the corresponding discussion forum. This decision took into consideration that each forum is different according to factors such as size and characteristics of member base, area of interest, quality and length of posts, and level of forum activity. For example, a high WOM level of a new ski in TGR may only represent a low or medium WOM level in a much larger and active forum such as Newschoolers.com. The study differentiated between three categories of PI and WOM: low, medium, and high. The cutoffs for these categories were calculated using the 25th and 75th percentiles in each individual discussion forum.

The market performance of the new ski models was determined on the basis of first year sales figures (where available) and qualitative performance data obtained for each new ski model. Researchers had to read sales figures in relation to the ski manufacturers’ largely different sizes of total sales. For example, a sales volume for one ski, which may be considered disappointing for a big player, such as Atomic, may constitute a significant amount of total sales for a small and highly specialized producer such as Armada. Additionally, the skis analyzed in this study each belong to different usage segments, namely freestyle, freeride, and all-mountain. These usage segments vary in terms of target group size. A comparison of absolute sales figures to compare market performance would not make sense. A categorical variable with two levels: “market success” and “market failure” served to classify new ski model performance depending on sales figures in view of each ski manufacturer’s total sales volume as well as based on the product success perceived by top executives a year after market launch.
3.3. Results
Almost 70% of all coded opinions were either positive or very positive. In comparison, only 244 or 18.06% of all opinions were negative. From the content of online discussion one could conclude that the active members of each forum are experts in a specific style of skiing and in related product categories. (Tab. 4). Despite the fact that members of EpicSki and TGR, for example, do discuss some freestyle skis forum discussions concerning products outside of a forum’s product category focus showed low levels of product interest and word-of-mouth as well as a low number of new product opinions.

<table>
<thead>
<tr>
<th>Online Skiing Forum</th>
<th>Product category focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>EpicSki</td>
<td>All-mountain and race skis</td>
</tr>
<tr>
<td>TGR</td>
<td>Freeride/Powder specific skis</td>
</tr>
<tr>
<td>Newschoolers.com</td>
<td>Freestyle skis</td>
</tr>
</tbody>
</table>

Table 4: Online skiing forums’ area of expertise

<table>
<thead>
<tr>
<th>Ski Model</th>
<th>Forum</th>
<th>PI Level (no of views)</th>
<th>WoM Level</th>
<th>Forum Opinion Before ML</th>
<th>Sales Data</th>
<th>Qualitative Data</th>
<th>Market Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armada ANT</td>
<td>NS</td>
<td>3214 Low</td>
<td></td>
<td>S Pos. 100.00%</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armada ANT</td>
<td>TGR</td>
<td>14140 Medium</td>
<td></td>
<td>S Pos. 100.00%</td>
<td>250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nordica Hot Rod Top Fuel</td>
<td>EpicSki</td>
<td>28189 High</td>
<td></td>
<td>S Pos. 91.18% Neg. 2.94% Neu. 5.88% n.a.</td>
<td>Best selling Nordica ski for 2005-06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LINE Invader</td>
<td>NS</td>
<td>n.a.</td>
<td></td>
<td>S Pos. 72.50% Neg. 10.83% Neu. 16.67% n.a.</td>
<td>100% retail store sell out</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salomon 1080 Foil</td>
<td>EpicSki</td>
<td>6276 Low</td>
<td></td>
<td>US Pos. 57.14% Neg. 42.86% n.a.</td>
<td>After initial consumer confusion, ski sold well in retail</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salomon 1080 Foil</td>
<td>NS</td>
<td>n.a.</td>
<td></td>
<td>S Pos. 72.22% Neg. 22.22% Neu. 5.56% n.a.</td>
<td>See above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salomon 1080 Foil</td>
<td>TGR</td>
<td>2300 Medium</td>
<td></td>
<td>US Neg. 66.67% Neu. 33.33% n.a.</td>
<td>See above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Völk Unlimited AC4</td>
<td>EpicSki</td>
<td>18415 Medium</td>
<td></td>
<td>S Pos. 85.71% Neu.14.29%</td>
<td>4000 Sold out in Dec 2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Völk Mantra</td>
<td>EpicSki</td>
<td>4803 High</td>
<td></td>
<td>S Pos. 78.26% Neg. 6.52% Neu. 15.22%</td>
<td>3000 Sold out in Dec 2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Völk Mantra</td>
<td>TGR</td>
<td>25281 High</td>
<td></td>
<td>S Pos. 69.70% Neg. 6.06% Neu. 24.24%</td>
<td>See above</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Successful new ski models

Product interest and WOM levels provided an indication of whether a ski could be considered a success or a failure within an online skiing forum. Tabs. 5, 6, and 7 give an overview of the product interest levels, the word-of-mouth levels and forum opinions concerning the new ski models analyzed in the discussion forums. The tables also provide sales and qualitative data concerning market success in the first year after product launch. Skis with successful, unsuccessful, and unknown market performance where grouped together to form three distinct tables. The column “Forum Opinion before ML” indicates the shares of opinion valences before market launch followed by the categorized overall valence of product discussions, either successful (S) or unsuccessful (US). Overall, the ski innovations analyzed in this study included six successful and two unsuccessful launches of new products. In three cases, the market performance could not be determined due to the fact that neither quantitative nor qualitative data was attainable. With one exception, all ski models characterized by low levels of PI and/or WOM were unable to fulfill the cut-off levels for classification as successful inside the discussion forum.
The number of successful ski innovations in the sample appears quite high considering that new consumer goods failure rates range between 36 and 86 percent (Franke 1998). An explanation for this finding is that WOM communication regarding new products in online discussion forums tends to be positive. Unattractive new products often do not even attract forum members’ attention.

Two approaches served for determining the forecasting potential of new ski discussions in online forums. A visual comparison of the internal evaluation of new ski models before market launch with the actual success or failure of the ski models in the marketplace (Tabs. 5 and 6) showed that in 12 out of a total of 14 cases, the valence of online forum discussions before market launch corresponded with the market performance of the analyzed new ski models. The valence of discussions concerning the Salomon 1080 Foil in both Epicski and TGR did not allow correctly forecasting market performance. The Salomon 1080 Foil was a freestyle ski designed for skiing park and pipe terrain. This ski only attracted low to medium levels of PI and WOM in Epicski and TGR together with large percentages of negative opinions and relatively low percentages of positive opinions. Both discussion forums do not attract experts in freestyle skiing. The expert forum in freestyle skis, Newschoolers.com, evaluated the Salomon 1080 Foil positively before market launch.

A logistics regression analysis served to quantitatively assess the influence of product interest (number of views), word-of-mouth (number of threads and number of replies) and valence of opinions (positive, neutral, negative) on market performance (see Tab. 8). Due to the small amount of cases (n=14), the potential level of analysis was constrained to single binomial regressions for every independent variable (Hair et al. 2010). Normalization of the volume measures before running the regressions allowed comparing the levels of product interest and word-of-mouth across the three discussion forums despite their different sizes.

The results show that positive WOM in the discussion forums is strongly associated with new ski market per-
performance. The change in the log-Likelihood function as well as the highly significant chi-square test (p < .01) indicate an acceptable overall model fit. Cox & Snell’s R² (.502) and the Nagelkerke R² (.682) show that the model accounts for a substantial amount of the variation in the market performance measure. The model is able to correctly classify 84.6% of all cases, which is significantly above the proportional chance criterion of 61.5%. All other independent variables had no significant influence on the ability to predict the market performance of new skis.

5. Discussion

Accurately predicting the performance of a new product in its stage of market introduction is of major importance to the profitable allocation of company resources. This study aimed to determine the forecasting potential of observing prelaunch word-of-mouth in online discussion forums for market performance of new products and the impact of volume and valence of prelaunch new product discussions on that performance. The forecasts derived from the observation of prelaunch discussions on ten new ski models in three skiing forums showed impressive accuracy. The valence of published opinions before launch was significantly related to the market performance of the new products.

According to the literature, the volume and valence of online word-of-mouth in review sites concerning new motion pictures, books and music albums are related to the early market success of these products. Findings about which of the two factors is more influential have been inconsistent so far (Chen et al. 2004; Chevalier/Mayzlin 2006; Chintagunta/Gopinath/Venkataraman 2010; Dhar/Chang 2009; Duan/Ga/Whinston 2008; Liu 2006). From the observation of online discussions of new ski models in three online discussion forums before product launch the valence of word-of-mouth emerged as an impressively accurate predictor of market performance. The volume of prelaunch discussions had some but no significant association with market success. This difference to previous research findings may be due to several facts. In the research presented here all word-of-mouth data was gathered well before product launch. In most of the formerly published studies, data came from previews close to the launch or from shortly after product launch. The motives for participation in and use of review sites just around product launch may be different to the participation and use of discussion forums well ahead of product availability on the market. When the volume of reviews of a movie, a book or a music album is high this may be a trigger for potential consumers to go see the movie, buy the book, listen to the album at a store in order to be able to participate in online and offline discussions. Buyers can form their own evaluations after consumption at relatively low risk. Skis are substantially higher priced than albums, books and movies. Most consumers cannot easily try them at low risk. Therefore, the valence of product information available from expert discussions online seems to be of greater importance.

According to the literature (Arndt 1967; Chevalier/Mayzlin 2006), published negative opinions have a stronger effect on other consumers’ opinions and behaviour than positive opinions. In the presented research no such effect could be found. Positive prelaunch opinions were related more strongly to market success than negative opinions. We can only speculate about the reasons for that finding. One reason might be that discussion forums in general contain more positive than negative evaluations (Godes et al. 2005). Active participants can become rather enthusiastic. These strong emotions may spill over to other participants and even “tourists”. On the other hand, the researched discussion forums contain a substantial number of highly respected experts in the specific field of interest of forum participants. In many cases the opinions of these experts concerning the new ski models were positive. The impact of these opinions on the overall valence of the discussions may have impacted the resulting behaviour.

The example of the Salomon 1080 Foil model shows how important it is to understand the area of expertise of each online forum in the attempt of forecasting new product market performance. Accurate forecasts can only be made based on content from online forums consisting of members who are highly involved in the specific product category and deeply understand the usage of the new product (Karimouchina 2011). Careful observation of discussions is needed to precisely determine the area of expertise of online discussion forums. Marketing managers and researchers should use discussion volume (that is, PI and WOM levels concerning the specific kind of product) as indicators of whether an online forum represents an adequate expert group in order to correctly forecast the market performance of a new product.

A comparison of gaining insights into consumer prelaunch word-of-mouth concerning new products in online discussion forums to traditional market research methods reveals important advantages. Studying product evaluations of consumers in online discussion forums is perfectly unobtrusive. Data collection does not place consumers in artificial environments, as is necessary with concept tests, conjoint analysis and prediction markets.

### Table 8: Correlation Coefficients of Word-of-Mouth Volume and Valence with Market Performance

<table>
<thead>
<tr>
<th>Kendall’s tau_b</th>
<th>Views_</th>
<th>Threads</th>
<th>Replies_</th>
<th>Positive</th>
<th>Negative</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>0.202</td>
<td>0.407</td>
<td>0.157</td>
<td>0.646(**)</td>
<td>-0.423</td>
<td>-0.333</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed)**
Compared to product-concept testing the observation of consumer discussion forums neither suffers from demand artefacts that are difficult to avoid in surveys, nor from biases introduced by the laboratory environment of experiments. The observation of online discussion forums does not measure individual preferences in a social void. Word-of-mouth can be observed as well as participants’ imitation of the behaviour of relevant others, or the impact of marketing activities, such as trade shows or events, on the volume and valence of new product-related discussions. Online forums offer unrestricted access to relevant consumer data, allowing time and money saving desktop research. Members’ opinions are accessible 24 hours seven days a week long before product market launch.

At the current state of online discussion forum research the potential market share of a new product cannot be determined. This disadvantage compared to conjoint analysis is partly compensated by the method’s much broader field of application. Conjoint studies are rather limited to products where the total benefit a customer receives can be reasonably defined as a linear combination of the new product’s attribute utilities. Online pre-launch discussions occur for most kinds of products, independent of the predominant manner of consumer purchase decision making, in all industries where insiders, freaks and highly experienced consumers have opportunities to receive information concerning new products before launch or even may try these products before market launch. Currently, the measurement of volume and valence of word-of-mouth in discussion forums before new product launch does not provide sales figure estimates, as do diffusion models. But when observing online discussion forums word-of-mouth effects, the level of imitation, price sensitivity and risk avoidance, the impact of social signals and network externalities need not to be estimated. They can be directly observed as forum members communally discuss experiences, evaluate product attributes and influence each other in their product opinions. There is no need for assuming that the members of a social network are fully connected or not. Network analysis tools allow tracking the emerging structure of communicative interactions in online discussion forums and measuring the impact of that structure on the new product evaluation process (Freeman 2004). A combination of pre-launch discussion forum observation and the estimation of diffusion models after launch may be an attractive avenue for future research.

Compared to prediction markets, at the current state of research the observation of online new product discussions before market launch has the disadvantage of not being able to predict the size of new product market shares. On the other hand, the tested method does not need to unrealistically assume independent sources of consumer knowledge. Highly active discussion participants use their own pre-launch product experiences enabled by marketing activities of new product producers as particular information sources. Others base their contributions on the information given in the forum, and a majority of participants silently observe the discussions as simple “tourists” who gain the information they need for their own decisions. The over-utilization of conspicuous information by consumers compared to information that is more difficult to obtain and to use, which may bias the results of prediction markets, does not appear to be a problem for forecasting new product performance based on the observation of discussion forums because no assumptions are made about the rationality of discussants’ behaviour. All participants in online discussion forums are self-motivated. Instead of monetary goals, like in prediction markets, genuine interest in the product, in related activities or in other participants drives contributors to be more or less active. The different mental models of participating consumers can be derived from their posts. No potentially interesting information is lost through a normalizing price metric.

In conclusion, the observation of prelaunch discussions in online forums is a promising method to forecast success or failure of products that can be exhibited to and tried out by expert customers before product launch who are interested in sharing their information and experiences with other customers on the Internet.

6. Limitations

Despite the impressive results of this research in terms of forecasting accuracy, it must be noted that the WOM and PI levels in the three discussion forums were determined only concerning ten selected new ski models. The study excluded other popular or unpopular skis discussed in skiing-related online forums. The potential interaction of word-of-mouth volume and valence could not be analyzed because the limited number of observed discussions did not produce a sufficiently large data set to apply moderated regression. An inclusion of all skis discussed within online skiing-related forums seems quite impossible, however, due to the sheer number of discussions.

The measure of new product market success may appear rather weak. Ski producers do not publish sales figures for individual models. Even at a distance in time producers are reluctant to disclose sales figures to researchers. The approach taken here is only a walking stick. For lack of better solutions, information gathered from interviews with top managers responsible for sales and major retailers should provide usable approximations of new product market performance. Future research will have to focus on other product categories where individual product sales data are publicly available.

The netnographic method used for data gathering is not without limitations. These limitations include the need for researcher interpretive skill and problems in generalizing results to consumer groups other than the online forum sample. Rigorous application of rules for organizing, coding and interpreting qualitative data can limit potential selection and interpretation bias. Representativeness of discussion forum participants does not seem to be
an issue. Research on different roles in such gatherings has shown that enthusiastic, actively involved, and sophisticated insiders demonstrate lead user characteristics (von Hippel 1986) and are highly respected within the community (Kozinets 1999). They have a strong influence on the opinions of other active discussants and on decision making of passive “tourists”. In the presented study, the great accuracy of market performance forecasts based on forum discussion data are an indicator that potentially restricted representativeness of discussion forum members did not have a negative impact. Due to the careful selection of the ten new ski models analyzed, which included varying degrees of forum discussion levels, the study builds in fact on a valid basis of online ski forum discussion behaviour.

The findings of the presented study may not be generalizable to products, which have only little experience character and are not consumed or used in public. Such products may not arouse enough interest to become subject to intensive discussions of active consumers and visits of a great number of rather passive consumers on the Internet. In light of the many discussion forums to be found on the Internet, which just concern such products, this issue seems to need further inquiry.

7. Conclusions

The presented research aimed to find out if the volume and valence of discussions in specialized online forums before new product launch may be used as reliable predictors for the success or failure of consumer products in the introduction stage of market diffusion. Results show that analyzing new product discussions in online forums focusing on the product category of interest and its usage can provide marketing managers with a valuable tool to forecast the future success or failure of a new product before launch. Compared to widely used concept test surveys the measurement of volume and valence of new product prelaunch discussions results in more accurate forecasts. Compared to diffusion models extensively discussed in the literature the disadvantage of not providing any precise sales forecasts is largely compensated by timely performance predictions based on observed instead of assumed or estimated behaviour. Compared to conjoint analysis and prediction markets the disadvantage of not being able to forecast market shares is compensated by the broader field of applicability and less risk of taking unrealistic assumptions.

References


Schlüsselbegriffe
Markterfolg neuer Produkte, Erfolgsprophose, Online-Diskussionsforum, Mundpropaganda

Keywords
New product market performance, forecast, online discussion forum, word-of-mouth